DOCUMENT RESUME

ED 032 542 AL 002 109

By-Adams, Edward N.; Rosenbaum, Peter S.

DLI-IBM Joint Feasibility Study in Computer-Assisted Foreign Language Instruction. Final Report.

Defense Language Inst., Monterey, Calif. West Coast Branch.; International Business Machines Corp., Yorktown Heights, N.Y. Thomas J. Watson Research Center.

Spons Agency-Defense Language Institute, Washington, D.C.

Pub Date [69]

Note-304p.

EDRS Price MF-$1.25 HC-$15.30

Descriptors-*Computer Assisted Instruction, *Language Instruction, *Military Training, Programed Instruction, Program Evaluation, Russian, *Second Language Learning, Teaching Methods

This document is the final report on a study of the use of computer assisted instruction (CAI). The objective of the study was to evaluate the potential applicability and usefulness of CAI in the instructional environment of the Defense Language Institute (DLI). The operational phases of the study were implemented in the Russian Aural Comprehension course at the DLI facility in Monterey, California (DLIWC). The study proved that (1) the instructional programming system was satisfactory as a prototype technology; (2) the instructional staff of DLIWC under DLI technical direction achieved good efficiency of program preparation; (3) operational integration of CAI into the DLI course and school was smooth; (4) CAI training was effective: students trained by CAI learned about as well as students in the normal Aural Comprehension course; and (5) CAI measures of student learning constituted highly reliable predictors of performance of trainees on DLI tests. It was recommended that a pilot CAI operation to develop CAI as a technique for future military training language instruction should be established at an appropriate time and place. It should develop improved CAI instructional techniques, the use of CAI as a language testing medium, instruction rates adjusted to student ability, and computer methods for improved evaluation and control. (Authors/DO)

ED0 32542

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

DLI-IBM JOINT FEASIBILITY STUDY IN COMPUTER-ASSISTED FOREIGN LANGUAGE INSTRUCTION

Principal Investigator: Peter S. Rosenbaum

FINAL REPORT

by

Edward N. Adams and Peter S. Rosenbaum

International Business Machines Corporation
Thomas J. Watson Research Center
Yorktown Heights, New York 10598

Contract No.: DAHC15-69-C-0145

Period of Performance: 26 October 1968 - 30 June 1969

Prepared For

Director, Defense Language Institute
Headquarters, Defense Language Institute
U. S. Naval Station (Anacostia Annex)
Washington, D. C. 20390

AL 002 109

DLI-IBM JOINT FEASIBILITY STUDY IN COMPUTER-ASSISTED FOREIGN LANGUAGE INSTRUCTION

Principal Investigator: Peter S. Rosenbaum

FINAL REPORT

by

Edward N. Adams and Peter S. Rosenbaum

International Business Machines Corporation
Thomas J. Watson Research Center
Yorktown Heights, New York 10598

Contract No.: DAHC15-69-C-0145

Period of Performance: 26 October 1968 - 30 June 1969

Prepared For

Director, Defense Language Institute
Headquarters, Defense Language Institute
U. S. Naval Station (Anacostia Annex)
Washington, D. C. 20390

ABSTRACT

An evaluation was made for the Director, Defense Language Institute of computer-assisted foreign language instruction. Educational effectiveness and operational practicality of CAI in a government language school were tested under the supervision of the Research and Standards Division HQ,DLI. The operational phases of the study were implemented in the Russian Aural Comprehension course at the Defense Language Institute facility in Monterey, California (DLIWC).

Specific findings:

The instructional programming system was satisfactory as a prototype technology; the instructional staff of DLIWC under DLI technical direction achieved good efficiency of program preparation.

Operational integration of CAI into the DLI course and school was smooth.

CAI training was effective; students trained by CAI learned about as well as students in the normal AC course.

CAI measures of student learning constituted highly reliable predictors of performance of trainees on DLI tests.

General Conclusions:

Conversational CAI has good potential for DLI as an instructional medium, as a testing medium, and as a component of a management information system. Its use could ultimately lead to qualitative improvements in some areas of instruction, to more efficient use of instructional staff, and to major savings in instructional time.

Within a suitable systems context properly selected DLI instructional staff with no special computer training can successfully prepare CAI instructional materials and can integrate CAI into DLI language courses.

Recommendation:

A pilot CAI operation to develop CAI as a technique for future military training language instruction should be established at an appropriate time and place. It should develop

- improved CAI instructional techniques
- the use of CAI as a language testing medium
- rate of instruction adjusted to student ability
- computer methods for improved evaluation and control of language courses.

ACKNOWLEDGEMENT

In presenting this project report from IBM to the Director, Defense Language Institute, I want to emphasize that the work reported is to a great extent that of the DLI staff itself. Clearly the project could not have been designed and implemented on such a tight time scale without the committed efforts of those involved, whether DLI or IBM staff.

Special acknowledgement is made to the Project Officer, Dr. Norman Levin, HQ,DLI and the following DLI WC personnel, Mr. Alex Albov, Director, East Slavic Division, the late Mr. Andrew Christiansen, of the East Slavic Division, and Mr. Shigeya Kihara, Director, Research and Development, for contributions to the project design and cooperation and help throughout the project. Special acknowledgement is also made of the contribution of Mr. Serge Issakow of the East Slavic Division, who, upon the death of Mr. Christiansen three days prior to the operational commencement of the project, stepped in to direct and lead the teams that produced and taught the experimental course. The materials of instruction were the work of a team of staff members of the East Slavic Division, DLIWC, Mr. Viktor Volmensky, Mr. Nicholas Marchenko, Mr. Michael Chordas, Mrs. Kyra Markov, and their secretarial assistant Mrs. Elena Nicolayef, who got a quality job done ahead of schedule.

Acknowledgement is also made to Professor Richard Atkinson of Stanford University and Professor H. William Morrison of Stony Brook University who served as valued consultants on the experimental design.

I wish also to acknowledge the contribution of the IBM staff involved, Mr. John T. Bednarik, IBM FSD for his contribution in coding the control program and managing technical operations on-site at Monterey; Mrs. Barbara Owens, IBM, SDD, Mr. Paul S. Cohen, Miss Janet Futvoye, Miss Susan B. Tevlin, IBM Research for help with various phases of program preparation and operation; Dr. John Emrick, IBM Research for his help with the experimental design and data analyses; Mrs. Jacqueline A. Lapolla and Mrs. Gloria C. Kneen, for the preparation of the final report; and Dr. Edward N. Adams, IBM Research, Director, Computers In Education, for his contributions to the design and analysis of the experiment and to this final report.

Peter S. Rosenbaum, IBM Research
Principal Investigator

TABLE OF CONTENTS

Page

INTRODUCTION 1

BACKGROUND 5

Phase I--Planning and Course Preparation 9

Phase II--CAI Operations 11
    Subject Selection 11
    Operational Statistics 15
    Operational Adjustments to Scoring Parameters 17
    Student Questionnaire 20
    Data Collection 21

Phase III--Evaluation 23
    Evaluation of Effectiveness 23
    Subsidiary Evaluation 42
    Discussion of Results: operational questions 57
    Effort Analysis and Cost Projection 59

Directions for Future Work 74

Table of Contents of Appendices 82

Introduction

This document is the final report on a study of the use of computer assisted instruction (CAI) for foreign language instruction conducted for the Director, Defense Language Institute, (DLI), Washington, D. C., by the instructional staff of DLIWC, Monterey, and the International Business Machines Corporation (IBM). The objective of the study was to evaluate the potential applicability and usefulness of CAI in the instructional environment of the Defense Language Institute and the government.

The evaluation was conducted under the supervision of the Division of Research and Standards, Headquarters DLI, Washington, D. C. Five questions were proposed for exploration:

1. Would CAI-audio exercises be of significant instructional value for training DLI military language trainees?
2. Would productive-type written exercises improve oral production?
3. Could the favorable distribution of skills often found in CAI experiments (i.e., relatively better learning by weaker students) be achieved in DLI?
4. Would CAI performance scores prove to have value as student achievement tests in relation to the DLI instructional objectives?

5. How well, qualitatively and quantitatively, could the staff of a DLI school deal with the operational problems involved in producing CAI materials and adapting school procedures to make use of CAI methodology?

Implementation of the objectives of the study involved

1) preparation by DLI staff of a substantial amount of CAI materials of instruction
2) the operation under DLI direction of an experimental AC course in which CAI was used as a partial replacement for conventional AC instruction,
3) joint evaluation by HQ,DLI and IBM of the instructional effectiveness of the CAI component of the course.

Operational phases of the project were conducted at DLIWC in Monterey, California. The first phase began on September 16, 1968 with the authoring of instructional materials. The second phase, the operation of the experimental section of the 4-68 AC course, began on December 2, 1968 and ended March 28, 1969. The third phase of the study, a formal evaluation, began on March 28, 1969, and concluded on May 30, 1969.

The study is the continuation of an evaluation activity of the R and S Division of HQ,DLI and the Academic Advisor DLI, who for several years have monitored research in CAI language instruction, including that done with the experimental IBM 7010-1440 CAI computer system at Yorktown Heights, New York. Implementation of the study was based on the availability of special equipment and programs existing only on the experimental 7010-1440 system. In order to exploit these programs and facilities for the operational work at Monterey computer terminals at Monterey were connected to the central computer via ten cross country telephone lines.

Access to communications equipment was possible only during a limited daily time period, and access to both communications and computing equipment only in the period prior to April 1, 1969; these time constraints were crucial in determining the timing of the phases of this project and the scope of the experimental work.

This report is organized in eight sections

1) Background
2) Phase I--Planning and Course Preparation
3) Phase II--CAI Operations
4) Phase III--Evaluation
    I Evaluation of Effectiveness
    II Subsidiary Evaluations and Analysis
    III Discussion of Results

IV Effort Analysis and Cost Projections
V Directions for Future Work

Origin, Nature, and Purposes of the Experiment

BACKGROUND

For some time prior to the initiation of the DLI-IBM study, IBM research had been studying methods of improving the process of language instruction. In a research study conducted at Stony Brook (SUNY), the use of CAI as an adjunct to classroom instruction led to (a) significantly increased learning and (b) a more desirable distribution of skills in that weaker students were strengthened relatively more than stronger students.

It was thought that the positive results from use of CAI were obtained not so much because a computer was used, but because the particular computer program implemented certain basic requirements of good foreign language instruction. Good foreign language instruction is believed to involve extensive student practice of complex learning tasks involving linguistic behavior; in the performance of these tasks the learner should be continuously supervised and should receive frequent, individualized correction of errors to forestall the development of poor language habits; moreover, the learner should continue practice until he develops the desired level of language proficiency.

The CAI program developed for use in the Stony Brook experiment met the above requirements. It provided

1) generative exercises requiring the student to produce complete utterances in the foreign language,
2) individualized correction of errors
3) a supervisory system which allocated the amount of work a student had to do on a particular topic in accordance with continuously computed measures of proficiency.

An important feature of the CAI foreign language programs was their tutorial character. The tutorial features of an ideal learning situation for reading and writing are difficult to achieve in a normal classroom, where any particular student can spend only a small percent of his time in teacher supervised practice. The CAI foreign language programs provided for such supervised practice of reading and writing in the CAI laboratory where the students received tutorial-remedial comment on their work from the computer.

Previous IBM experiments with CAI language instruction in the university had been confined to elementary phases of language instruction, and had given most emphasis to the skills of reading and writing. From the IBM view DLI represented a near ideal setting in which to carry out research on more advanced foreign language instruction. DLI has a suitable student population; can provide and maintain very effective control of operational variables; has a large number of class cycles started and completed each year; has a teaching faculty well qualified to prepare linguistic materials of high quality; has audio-lingual training objectives about which to develop and evaluate audio based CAI exercises which have not yet been extensively evaluated in a university setting.

The Research and Standards Division of HQ,DLI began evaluating the potential of the IBM CAI language program in 1967 while the first set of evaluation experiments were in progress at Stony Brook University. This evalu-ation began with site visits by the Academic Advisor, DLI, and staff members of the R and S Division. These visits were followed in 1968 by a technical briefing of the Director, DLI and members of his staff.

In late 1968 Headquarters, DLI supervised a practical evaluation in which faculty members from the DLI East Coast branch, Anacostia, (DLIEC) prepared CAI exercises in three different languages as a probe of the practical problems of a number of questions of operational feasibility. These experiences indicated that an operational evaluation in a DLI facility under

DLI managerial and operational direction would be feasible with the expenditure of moderate resources of time, money, and professional effort.

Both IBM and HQ,DLI were aware that if CAI were effective in relation to DLI's training goals, it could provide special operational advantages for DLI:

1) Highly efficient use of training time.
2) Reduction of student failure.
3) Adaptability for quick modification of old courses, development of new courses, and build up of instructional facilities.
4) Increased flexibility to shift training resources from course to course as well as teacher skills.

For these reasons an experiment to evaluate the applicability of CAI language instruction in the DLI environment seemed very worthwhile,

Course Design and
Production

PHASE I-PLANNING AND COURSE PREPARATION

The advanced plans for the project were developed by the IBM Education group and HQ,DLI. The operational plans were developed jointly by HQ,DLI, the staff of the DLIWC branch, and the IBM Education group.

The advanced plan specified that CAI would be used for a substantial portion of instructional time, with periodic operational reviews to ensure that the DLI trainees should not fall behind DLI training goals.

Headquarters DLI decided that the operational experiment should be carried out in the Russian Aural Comprehension (AC) course at DLIWC. An AC course was chosen because its emphasis on comprehension rather than on oral production agreed with the nature of the audio-CAI methods to be used. Russian was chosen for several reasons, related both to convenience of experiment and the possibility of quick followup if good success was achieved.

After the advanced plan was agreed on, meetings were held at DLIWC to work out the operational plan. The operational plan included the following:

1) A plan for an experimental AC course making use of CAI.
2) Selection of exercise types, lesson designs, and computer scoring systems to be used.

3) A detailed plan of the course development phase.

4) A detailed plan for the operational phase.

5) Identification of managerial and professional staff.

6) A general plan of final evaluation.

The completed plan required that the lesson designs and control program be completed by September 16, 1968 and the instructional materials completed by December 3, 1968. The evaluation of the experimental AC course with CAI was to be made with the AC class that completed its third week at the end of November. The analysis of results was to be completed by May 30, 1969 and a final report by June 30, 1969.

A detailed account of the course design is attached as Appendix I, sections A and B. It includes an account of the design of the experimental AC course, the CAI control program, and the CAI exercises. The technical sections of the course production plan are in Appendix I, section C.

CAI Operations

PHASE II-CAI OPERATIONS

SUBJECT SELECTION

The experimental course was administered to the AC class beginning in November, 1968. It had been expected that this class would contain approximately ninety students, of which thirty would be selected for the experimental CAI group and a portion of the remainder as a reference or control group. Students in the CAI group were to be as closely matched as possible with the students in the non-CAI group on the basis of both the Army Language Aptitude Test (ALAT) scores and grades on the DLI 3-week test.

I. Initial Selection Procedure

- A. On December 2, 1968, thirty of eighty-two students in the November '68 Russian AC class were selected to form the experimental CAI section for the DLI project in computer assisted instruction. The selection procedure used was as follows:
  - 1. A scatter plot was prepared of the ALAT scores and the DLI academic grades at the end of the first three weeks of instruction at DLI (Appendix II, Fig. II-A-1). (The DLI three week grade is computed as

the average of the second week class grade, the third week class grades, the 3-week oral test, and the 3-week written test.)

2. By inspection the points on the scatter plot were marked off into thirty groups, each group having either two or three students. Both or all three members of a group were closely matched on both ALAT score and DLI grade. (Appendix II, Fig. II-A-2).
3. A die was rolled to determine which one of the students in each group was to become a member of the experimental section (Appendix II, Fig. II-A-3).
4. As a check on the similarity of composition of the two groups the distribution of students by ALAT and DLI grade was plotted for both the experimental and the reference sections. (Appendix II, Fig. II-A-4) The resulting distributions showed that there was close matching of the groups at all levels of skill.

B. Students were informed of the selection results on December 3, 1968 and CAI instruction began ~~at~~ 1400 on that day.

C. It was expected that during the early weeks there would be some dropouts from the experimental section for academic and possibly administrative reasons since several of the original student group were very weak. Accordingly a preselection of replacements was made for any students who dropped from the CAI group in the first few weeks. The CAI group was divided into quartiles and alternates were selected for each quartile from the original groups containing these members. On the assumption that dropouts for academic reasons from the top two quartiles would likely be minimal, only one alternate was selected from each of these quartiles. Matched groups in the lowest quartile all contained pairs rather than triples. Thus, alternates for the lowest quartile were selected from the third quartile as were the third quartile alternates. Four alternates were selected in the third and fourth quartiles, in anticipation of a higher drop rate in these quartiles.

II. Changes in Experimental Group

A. Shortly after the beginning of the operational phase, one of the CAI students requested transfer from CAI to non-CAI. This student, number 81 (see Appendix II, Fig. II-A-3), was switched with student 18, one of his matched group counterparts.

B. After the six-week test at the end of the first three weeks of operation, two students in the CAI group were dropped as failing. According to the initial plan, they were replaced by two students, numbers 59 and 63, originally placed on the CAI alternate list. After these adjustments some minor changes were made to the designated composition of the matched groups so that 30 matched groups remained after the reassignments.

The results of this readjustment are shown in Fig. II-A-5 (Appendix II). It can be seen from Fig. II-A-6 (Appendix II) that the two groups, CAI and non-CAI, retained a similarly well matched distribution with respect to the selection variables. It will be observed from Fig. II-A-6 that because of drops of non-CAI students at the end of the 6-week tests, the two lowest of the original matched groups (students 23, 21, 49, and 31) disappeared altogether. During the 12 weeks of operation after the six-week tests, no further changes were made in the composition of the CAI group except to drop failing students.

C. By the time of the 18-week test, which was administered during the week of March 23, four more students had been dropped from the class. (see Appendix II, Fig. II-A-7).

Operational Statistics

CAI operations commenced on December 3, 1968 and terminated on March 26, 1969. The original course development plan called for 127 hours of instruction organized into seventy-one two- or one-hour units (the latter for Fridays). Of this total, six hours of instruction were lost, four because of the administrative necessity of giving the MLA subtests during class time at the 12 and 18 week test points. An additional two hours were lost owing to adverse weather conditions in Yorktown Heights, which forced the closing of the Research Center where the IBM 7010-1440 computer is located. Miscellaneous problems caused the cancellation or shortening of a few classes for certain groups of students; total loss of CAI class time from such causes did not exceed an average of 1½ hours per student.

Computer equipment performed well and at no time was a student turned away because of equipment unavailability.

There was an initial problem with loss of service because of telecommunications failures. Telecommunications involved a connection to Yorktown Heights, New York through the IBM Dial Network via ten off-premise extensions of the IBM San Jose labs. Interruptions in communications, which arose largely from unknown events in the communications net, normally caused a "down" time for the terminal of only about two minutes. After the installation "shook down" these problems were not severe; a student could expect on average to experience less than one line failure a week.

OPERATIONAL ADJUSTMENTS TO INTERNAL SCORING PARAMETERS

As mentioned earlier the CAI program generates two sets of four weighted recency scores that are available to be used both for the flow of control in the CAI program and the administrative supervision of the students. The first set, the lesson recency scores, comprises indices of short term performance (one to two days) for grammar exercises, translation exercises, and aural comprehension. A corresponding set of four course recency scores for each class of exercises comprises similar indices of long term performance (1 to 3 weeks).

The range in time of these recency scores depends on weighting factors that can be under day-to-day control of the course administrator. Initially the weighting factors for the lesson and course recency scores were set at 2 and 9 respectively. Thus, the lesson recency score reflected most heavily student performance on the preceding three modules of particular classes of exercises and the long range course recency score reflected the student performance on the most recent ten CAI sessions. The lesson and course recency scores themselves were set initially at 80.

The course and lesson recency scores produced each day by CAI students at DLI were displayed daily. Monitoring of the long term recency scores during the first six weeks of the operational phase revealed that the distributions for grammar, translation, and dictation were gradually becoming systematically "flattened" with a preponderance of scores at the high end of the range (see Appendix II, Fig. II-C-1). Such a flattening meant that students of medium strength were generating scores nearly as high as the strongest students. Thus, on these three item types the CAI scoring system was failing to differentiate adequately among the performances of students in the upper half of the performance distribution, hence was not providing maximal information for purposes of controlling student work assignments.

It was decided at the end of the sixth week to adjust the item scoring parameters to increase their sensitivity to performance differences between students. The initial choice of parameters gave the student a great deal of opportunity to improve his score on retrial, so the flattening of the distribution could mean that we were being perhaps too lenient toward errors on the first attempt. On this reasoning the set of weights 3-3-4 were replaced with the set 9-1-0, which caused the item score to be determined 90% by the score on the first attempt, 10% by the score on the second.

No scoring change was made for aural comprehension type modules, since the distribution of scores on this type of exercise showed no such flattening. The new choice of scoring parameters was adhered to for the remainder of the course.

To assure that the readjustment of the long term score settled down to a stable value by the time of the 12 week test, the course recency weight was changed from 9 to 5, thus basing the computation of daily course recency scores on a six rather than a ten session basis.

One function of the recency scores is to provide criteria to the control program as to whether a student is allowed to skip supplementary modules, when he is thought to have had enough practice on the obligatory ones. However, it was recognized early in the course that DLI students were moving extremely rapidly and were generally more proficient than had been anticipated. Consequently, even the poorest student, after a few weeks, was capable of completing the entire lesson. Clearly we had not made enough material. For this reason, CAI program parameters were set during most of the course so that all students had to complete all items; only in the last few weeks were students with the very highest scores allowed to skip some supplementary work.

STUDENT QUESTIONNAIRE

In the last week of the course the students were asked to complete a questionnaire by means of which they could indicate their reactions to various features of the equipment and the program. They were also asked to make any comments they wished as to how the course might be improved. A summary of the students' judgments and a transcript of their suggestions are in Appendix III, Section H.

DATA COLLECTION

For purposes of analysis and evaluation the following data were collected:

Test Scores

1. Scores on ALAT and 6-, 12-, and 18-week tests on at least one hundred and fifty students in AC classes completed in the year prior to the completion of the experiment.
2. ALAT scores on all students in the November AC class.
3. 3-week test scores for the November AC class.
4. 6-week test scores for the November AC class.
5. 12-week test scores for the November AC class.
6. 18-week test scores for the November AC class.
7. 24-week test scores for the November AC class.
8. MLA Foreign Language Cooperative test scores for the November AC class (test to be administered during the week of 24 March 1969).

CAI Performance Indicies

9. First attempt scores--the average over the whole course of scores achieved by students on the first attempt on all items of a class:
   - (a) on all items taken
   - (b) on grammar drill items only (i.e., TST, ACST, ACT)

    (c) on translation items only (i.e., TRANS)
    (d) on dictation items only (i.e., DICTS, DICTNU, DICTNA)
    (e) on aural comprehension items only (i.e., ACS, ACD, IRN)

10. Module scores--the average over the course of all module scores of a class:
    (a) on all modules taken
    (b) on grammar drill modules only
    (c) on translation modules only
    (d) on dictation modules only
    (e) on aural comprehension modules only

11. Ten sesssion recency weighted scores taken at the time of the 6-, 12-, and 18-week tests:
    (a) for grammar drill only
    (b) for translation only
    (c) for dictation only
    (d) for aural comprehension only

12. Number of items completed (daily average computed at course end):
    (a) on all items
    (b) on grammar drill only
    (c) on translation only
    (d) on dictation only
    (e) on aural comprehension only

Results and Conclusions

PHASE III-EVALUATION

I. Evaluation of effectiveness

A. Student performance data

The student scores and grades are presented in Tables III-B of Appendix III. For purposes of easy visual comparison they are also plotted as graphs and presented in sections III-C and III-D of Appendix III; in these graphs the scores of non-CAI students are weighted according to the distribution of students among the original matched groups so that the total weight of non-CAI scores is the same as the total weight of CAI scores (as explained in the appendix).

The data summaries in tables III-E-9 and III-E-10 present for each of the criterion tests the mean grades or scores, their standard deviations, and the magnitudes of the various components of the overall standard deviation, each expressed in standard deviation form as a number of grade or score points such that the square of this score represents the variance associated with the variable. For example, the 3.4 points associated with treatment on the 18-week dictation test means the variance associated to treatment is the same as though there were an average grade difference of 3.4 points between the two groups.

The digest in Appendix III, Section F shows the means by quartile and by treatment of the various criterion scores. The adjusted scores are also plotted to provide perspective in comparison.

Time data

The time required for each student to complete his daily assignment was collected and the percentage of available time he used *on days he was present* was computed. These data are shown in Figure III-K-8.

12-week Academic Level

Scores on the MLA tests Form MA taken at 12 weeks were averaged to obtain the mean scores on each of the four test components. Comparing these mean scores to norms in the Educational Testing Service *Booklet of Norms* one finds that at the time of the 12-week test, scores of the "average DLI student" in the study had centile standings relative to those of high school students completing the third year of Russian study as follows:

Reading	42 centile
Writing	43 centile
Speaking	50 centile
Listening	49 centile

Centile Norms for first year college students on the MA were not available. It would appear, however, that the CAI centiles would have been somewhat above median for first year college.

18-week Academic Level

At the end of 18 weeks the means of scores on the four components of the Russian MB test showed that in relation to college students completing the second year of Russian, the "average DLI student" in this study stood at the following centiles:

Reading	40 centile
Writing	44 centile
Speaking	77 centile
Listening	68 centile

24-week Tests

The DLI 24-week tests grades were obtained for the purpose of followup examination. The raw data are presented in Appendix III-B. A graphical presentation of these data based on the surviving matched groups is given in in Appendix III-K. It is to be noted that two CAI students had dropped during the period after the 18-week test, neither of them for academic reasons. No further analysis of these grades was made.

B. Statistical analysis of performance data

The effectiveness of instruction in the experimental course was analyzed according to the experimental design, which is attached as Appendix III, Section A, Part One. The principal analysis was a treatment-by-levels analysis of variance to test for effects due to 1) method of instruction 2) ability grouping.

The criterion measures of performance were the DLI grades and scores on the MLA Foreign Language Cooperative Tests, both recorded at the time of the 12-week and the 18-week DLI tests. The independent predictors of ability were scores on the Army Language Aptitude Test (ALAT), taken before instruction began and the DLI grades at the end of the 3-week introductory portion of the course, which chiefly deals with phonetics and orthography.

The levels analysis includes a multiple regression to apportion the variance of scores among treatment, method, aptitude (ALAT), a second ability variable ("residual" 3-week test), and an "error variance" that is not correlated to any of the "predictor" variables. The second ability variable is a "reduced" or "residual" 3-week score from which the dependence on ALAT had been removed by a separate analysis of variance, so that the reduced 3-week test score could be considered as a measure of ability factors not measured by ALAT.

The principal analysis was done by quartiles, i.e., by four levels, and most of the results stated in this section are the result of this analysis. However, as the subsidiary analysis proceeded it became clear that the adjusted scores associated to the lowest ALAT quartile had high intra-cell error variances associated to them and had unexpectedly low correlations with other variables. The reason for such unsatisfactory behavior seemed to be that as students who dropped for academic reasons depleted the lowest quartile of its weakest students, the first quartile students who remained became indistinguishable on average from those in the next quartile. As a consequence, the predictive power of the ALAT was made to seem smaller than it really is and much of the variance that should be assigned to AIAT was assigned to another variable.

To verify the above interpretation a partial reanalysis by terciles, i.e., three levels, was made. With the students divided into terciles there were many more surving students in the lowest group and it is more reasonable to believe that the survivors in each level are now statistically representative. The results of tercile analysis show internal evidence of improved statistical quality in that the mean intra-cell error variances were reduced, and the amount of variance associated to the independent "predictor" or "blocking" variables was increased. Since the qualitative conclusions of the study were not altered in any important way by the improved analysis, new tables of adjusted scores, etc., corresponding to the tercile analysis were not prepared.

The adjusted quartile scores on the various tests are presented in Appendix III-F, and an analysis of the variance in these scores in Appendix III-E, where are tabulated separately those components of final scores or grades attributable to aptitude, treatment, interaction between aptitude and treatment, and ability independent of aptitude.

C. Levels of significance

Two factors have to be considered in estimating how important a particular influence or cause has been in determining how well students learned. One of these is how much of the differences among student grades is associated to that factor; the other is how large would be the "random" diffences expected under the circumstances of the experiment because of the action of factors that are not known or not under control. The random variance may be considered an "error" variance since it is a natural measure of the size of experimental errors under the given conditions. The "significance" of the variance due to a particular effect will be a function of how large is this variance relative to the "random" variance.

The present analyses employ the F-test of significance. The F value for a given source of variation is essentially the ratio of the variance associated to that source divided by the error variance. From the F value and certain other information one can estimate from F-test tables the probability that the experimentally observed differences would occur by chance when there is no real difference.

Two criterion levels of confidence are commonly quoted in behavioral work, the 5 percent level ($p < .05$) and the one percent level ($p < .01$). Thus the indication ($p < .05$) by a measured value indicates that a value so large would occur by chance in less than five percent of the cases if the "true" effect is negligibly small. Effects for which the probability of chance occurrence exceeds five percent are commonly referred to as "non-significant." We will use that terminology in this report.

D. Interpretation of results

Summary

1. Differences in learning attributable to the method of instruction (CAI vs non-CAI) were much smaller than differences attributable to aptitude and (scholastic) ability.

2. There were no significant differences in learning (CAI vs non-CAI) in the period 3 weeks to 12 weeks; there were apparent significant differences favoring the non-CAI students in the period 12-18 weeks as shown by performance on the tests at 18-weeks.

3. For these tests on which significant differences (CAI vs non-CAI) were observed the outcomes were biased to some extent by differences of secondary skills (e.g., typing vs writing with pencil) that are related to the method of learning but not necessarily to the ultimate instructional objectives.

4. When account is taken of the effect on test scores of differences in secondary skills, (CAI vs non-CAI), effectiveness of learning was very much the same for the two groups.

5. Differences in time required among CAI students were large; a CAI system with individualized scheduling could achieve significant time savings.

Detailed discussion of the scores

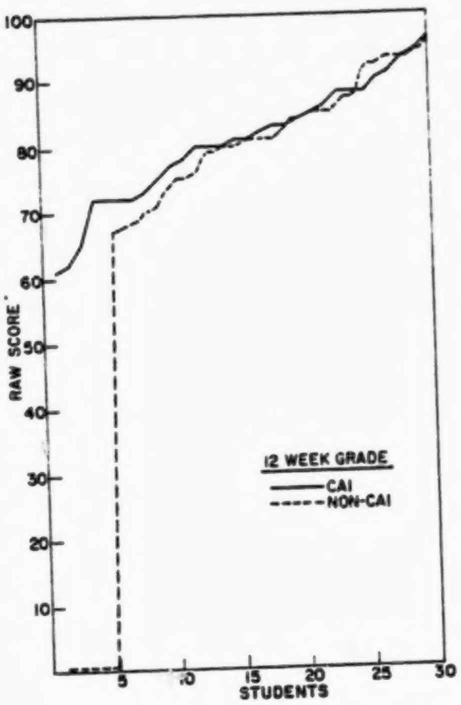
NOTE: the following discussion is a detailed examination of the performance scores; the general conclusions are given in the summary above.

Relative importance of treatment:

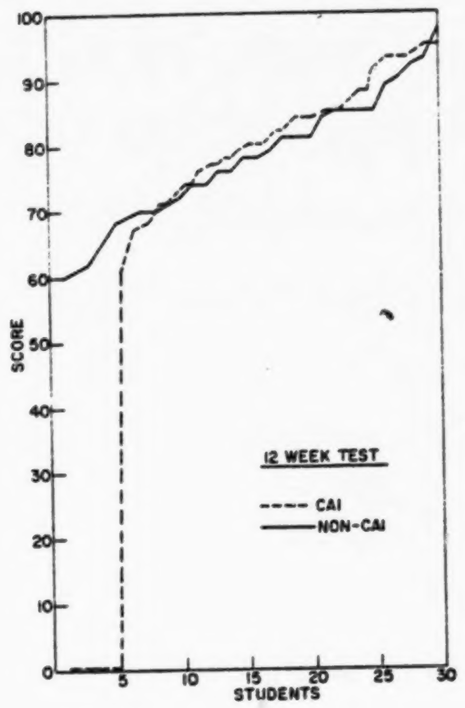
12-week Tests

There are no significant ($p < .05$) effects due to treatment (CAI vs non-CAI) on the 12-week test. (From the quartiles analysis the Aural Comprehension Sentences (ACS) test would seem to show a significant difference between CAI and non-CAI; however, this difference is judged to be spurious, since it does not appear in the "cleaner" terciles analysis and since subsequent aural comprehension tests show no differences of grade distribution between CAI and non-CAI students.)

Interaction between treatment and ALAT is generally unimportant. The data for the 12-week test and total 12-week grade are shown in Figures III-C-1 and III-C-2.



III-C-1



III-C-2

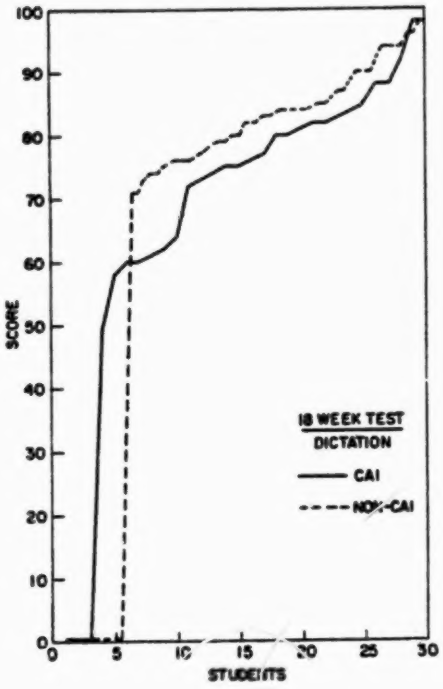
In these figures the individual scores have been plotted in order of rank with each score plotted against a length of abscissa proportional to its weight in its original matched group. The virtue of this type of plot is that it shows the relation between all scores in the two distributions including the drops; by contrast the use of the average grades of survivors as an index makes the performance of a group seem better if a larger number of students drop.

The figures show that the achievements of the two groups on the 12-week tests were closely similar when compared in relation to the original ability measures.

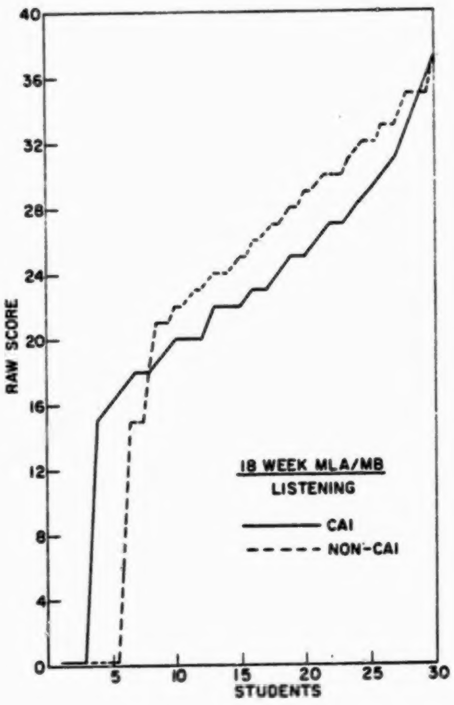
18-week Tests

Overall dependence of the 18-week DLI grades and MLA test scores on the 3-week test variable and the ALAT variable was similar to that on the 12-week test, with perhaps a slight lessening of the predictive power of the 3-week test. There were several significant differences attributable to treatment, on the MLA listening test ($p < .05$), on the DLI dictation ($p < .01$), grammar ($p < .05$) and overall test score ($p < .05$).

The most striking differences of performance are those on dictation and listening, which show a similar distributional pattern as can be seen from figures III-D-4 and III-D-12.



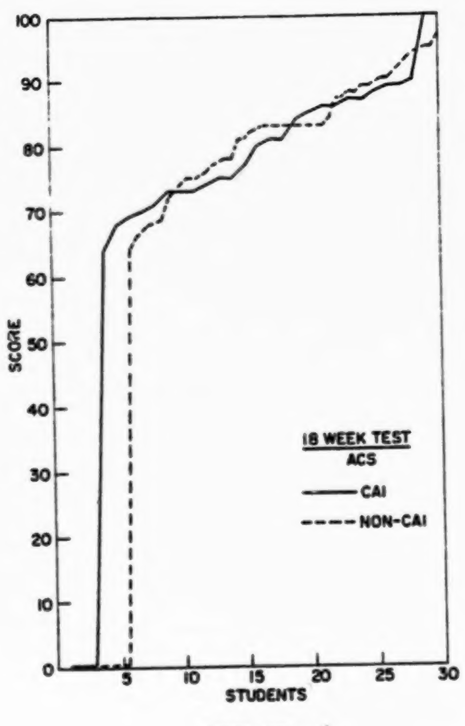
III-D-4



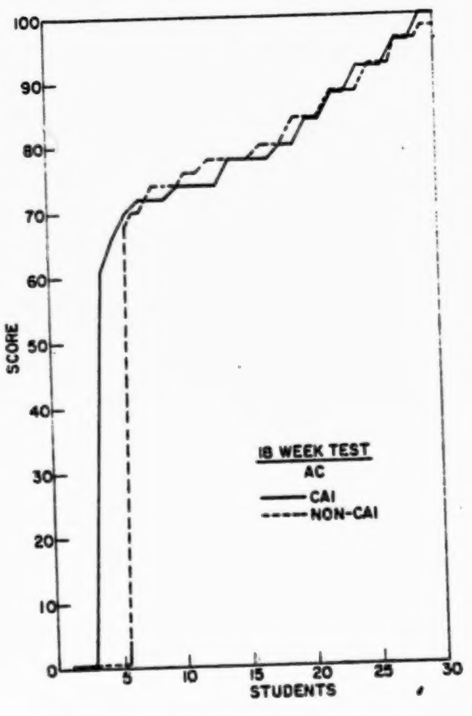
III-D-12

In both cases the CAI scores show a deficit of an approximately constant amount across most of the range of grades or scores, leaving little doubt that the CAI students are systematically poorer on some of the skills involved in these two tests.

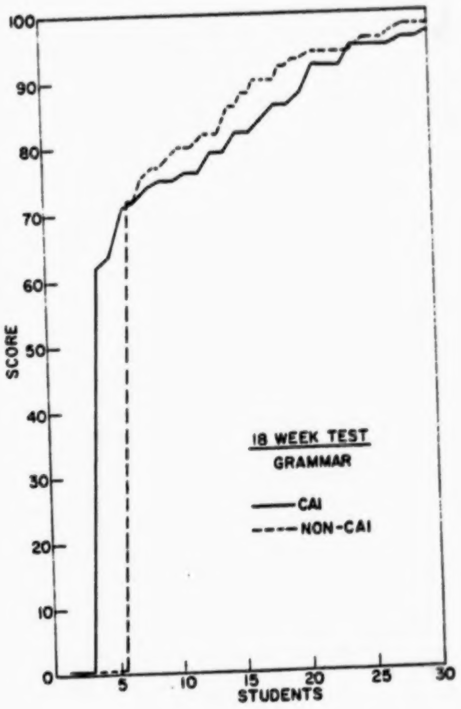
The nature of the skills deficit in these two tests is presumably not in the ability to hear accurately, since the CAI scores on Aural Comprehension of Sentences ACS and Aural Comprehension AC, Figures III-D-6 and III-D-7, show no signs whatever of a deficit of performance in these areas. A possible source of difficulty is in the relation between the testing task and the training tasks by which the students prepared, as will be discussed below.



III-D-6



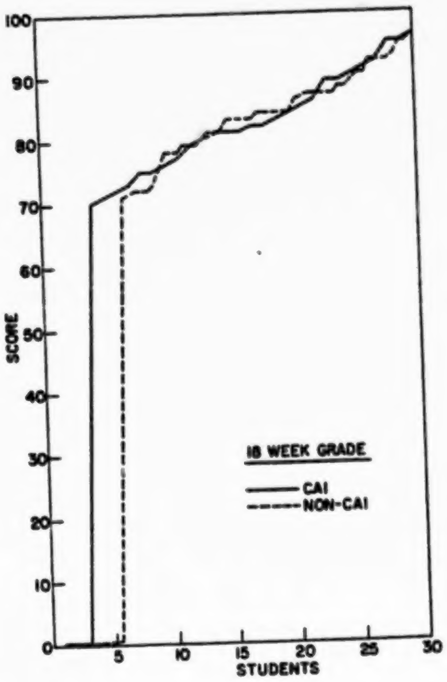
III-D-7



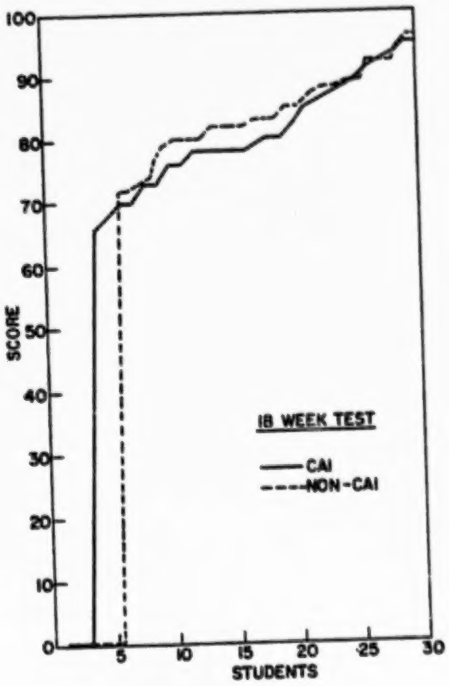
III-D-5

A systematic deficit in the 18-week grammar grades can be seen in Fig. III-D-5.

Figures III-D-1 and III-D-2 show the distributions of total 18-week grades and 18-week test scores. The total grade distribution shows, as did the 12-week grade, a close similarity of overall performance between the two groups. However, the 18-week test shows a statistically



III-D-1



III-D-2

significant deficit of performance for the CAI students deriving from the deficits in the grammar and dictation. By all other measures of performance the CAI students look even more similar to the non-CAI students than they did at 12 weeks.

Discussion

The evaluation tests were quite extensive; including the components of the DLI class grade, the DLI test, and the MLA battery, a total of twelve or thirteen tests were administered at both the 12-week and the 18-week test times. Apart from a few discrepancies, the overall picture is that the course using CAI was about as effective as the regular AC Russian course. As shown in Figure III-K-8 in obtaining these results the learners left unused about 22% of CAI training time. Since the final learning of CAI students is generally similar to that of non-CAI students these data imply a high level of effectiveness of this CAI program.

We will now examine those tests on which the CAI students did poorly to see whether they indicate a lack of effectiveness of CAI training. Let us first note that there appear to be no performance deficits in the first nine weeks of instruction (up to 12-week test). This result was not unexpected, since the prototypes for CAI exercises and lesson design had been developed for use with first year college students whose work is essentially at the level of the DLI students in the period preceding the 12-week test. The general effectiveness of CAI in this period confirms that in the DLI environment also, the CAI exercises are suitable to support learning of basic language mechanics.

Deficits of performance appeared on three of the 18-week tests, the DLI dictation, the MLA listening and the DLI grammar. Both the dictation and listening showed the same *distributional* pattern of deficit, an approximately uniform depression of all grades in the class. This pattern is so striking that it must indicate a real performance deficit of CAI trained students.

The first question to be asked is whether the deficit of dictation test performance corresponds to a real difference of *effectiveness* of *learning dictation*. It does not seem that the deficit can be one of aural comprehension, since the CAI students obtained fully satisfactory achievement on both of the aural comprehension tests. Similarly, it does not seem that dictation *as such* could be a problem, since the CAI students did very well on the CAI training tasks in dictation, averaging *as a group* 92.5% correct *on the first attempt for all dictated sentences in the course*.

It would be most surprising to find that CAI dictation exercises were ineffective. On its face dictation should be one of the most valid and most efficient of the CAI learning tasks. On the basis of faculty observation and student comments, we were led to suspect that the deficit on the DLI dictation test might be a deficit in secondary performance skills associated to the process of recording the dictation sentences.

The CAI training situation for dictation is quite different from the testing situation. In the CAI course the student received the dictated sentence on demand and recorded it at his own pace using a typewriter; in the test situation the sentence was presented under conditions of time pressure and was recorded with pencil and paper. It is clear that mediating processes that relate aural inputs to typing output are different from those that relate it to writing pencil and paper output; there should be further differences because of the absence of time pressure in the CAI training situation. We reasoned that the performance deficit of the CAI students in dictation most probably was due to deficits in skill in these hypothetical mediating processes.

Assuming these secondary skills are relatively minor, we would expect that transfer of responding skills to the pencil and paper recording mode under time constraints should occur rapidly after the CAI students were returned to the regular course taking daily dictation with pencil and paper. Consequently we examined the 24-week grades affected by dictation to look for evidence of change. Performance on dictation affects the 24-week class grade, since the daily dictation score is averaged into it, and also the 24-week test grade on dictation. The distribution of 24-week class grades shows that on average there was a perceptible deficit in the daily pencil and paper dictation of former CAI students during the period 19-24 weeks; the deficit in dictation is larger than the deficit in the 24-week class grade since the class grade includes several grades on skills for which the CAI students had no deficit. The poorer class grades of these students is in striking contrast to the good 18-week class grades they obtained when their daily dictation was still being done on a typewriter in the CAI mode; this supports the interpretations that the deficit is in the areas such as dictation which showed deficits on the 18-week tests.

In contrast to the 24-week class grade, the 24-week test in dictation shows little evidence for a general deficit; any deficit appears confined to at most a few students. Clearly at 24-week test time, after 6 weeks back in the normal course, the largest part of the transfer to the new testing conditions had occurred.

It is noteworthy, and favorable to our hypothesis that the deficits in the 24-week class grade, which reflect work done in the weeks 19 through 24 are not mirrored in the 24-week dictation test grade. This does not, of course, prove that the deficit in dictation is primarily related to secondary skills rather than to fundamental effectiveness of dictation learning, but it supports and is consistent with that hypothesis.

The 18-week grammar test, which also shows a deficit of performance for CAI students, is another case in which the testing situation encountered by CAI students is different from all the CAI training situations. So far as we can see from the internal CAI scores, CAI students did very well on the CAI grammar tasks, apparently better than on other CAI tasks. However, since the DLI grammar test scores correlate better with translation, dictation, and AC than they do with CAI grammar, we may suppose that the DLI grammar test may contain crucial elements of difficulty related to something other than the grammar patterns themselves.

We could not isolate those special elements of the grammar test that might be involved; however, we would expect as with the dictation recording skill that if the difficulty lies in some minor secondary skill it would quickly diminish when the CAI students are back in the regular AC course where such exercises are a normal part of the daily work. Indeed that is what happened as can be seen by comparing the 24-week grammar grades with the 18-week grammar grades. Thus as in the case of the dictation learning we are reinforced in the hypothesis that skills associated to the test situation rather than a true weakness in grammar knowledge are the most likely source of deficit on the 18-week grammar grade.

II. Subsidiary evaluations and analysis

Summary

No evidence was found that the written exercises used resulted in improved oral skills.

The distribution of skill of CAI students was similar to that of non-CAI students on essentially all types of language skill.

The overall quality of instructional materials was adequate. Certain of the exercise types were not fully suitable to the advanced material between the 12th and 18th week.

The test instruments available were not entirely suitable for evaluating the effectiveness of CAI learning in relation to ultimate DLI objectives.

Internal CAI scores were found to have high internal reliability and to be good predictors of success on DLI performance tests.

CAI internal scores showed high power for early prediction of student success.

Student reaction to CAI methods was positive and favorable to the operational use of CAI for a portion of instruction.

Detailed Discussion

The following section is an expanded discussion leading to the conclusions summarized above.

Value of Oral Exercise

Besides the primary questions relating to the effectiveness of CAI to support learning in the DLI environment, several subsidiary questions were to be investigated as opportunity permitted. One of these relates to possible contributions of productive-type written exercises to learning of oral production skills. Our design did not provide for direct experimentation with the oral production objective, but we looked for indirect evidence of a relation between oral production and CAI exercises in the form of correlation of CAI internal measures with the DLI grades on oral production or the MLA speaking test score. However, these correlations, though appreciable, as are all interskill correlations for these students, were actually smaller than other correlations between final skill and CAI performance. We conclude that the data do not especially support the value of written exercises for improving oral production.

An indication of the effectiveness of total oral training may be gotten by comparing the level of performance of the CAI and non-CAI students on the DLI oral and MLA speaking tests. This comparison shows that CAI performance was comparable to that of non-CAI students; however, it leaves unresolved how much of the relevant learning occurred in the portions of instruction that were different for the two groups.

Another matter of interest was to see whether CAI training resulted in an improved distribution of skill, one in which the weaker students do relatively better than in conventional classroom courses. No such effect was found in the present experiment; if anything the data show a very slight opposite trend: the top CAI quartile tends to be better than the top non-CAI quartile, the bottom CAI quartile poorer than the bottom non-CAI quartile.

We do not believe that this apparent trend in the data should be taken as a general indication of what to expect from CAI learning by military trainees. To bring about the desired distribution of skill we should provide:

(1) Enough practice material available so that the weakest students could achieve proficiency by working through it, and

(2) Enough practive time so each student can bring his skill up to the desired proficiency.

The conditions of our experiment did not allow us to give the weaker students the needed extra time and properly chosen supplementary work. In effect we gave the same tasks to all students, so our results merely show that the strong student learns the most from a given task.

Figure III-K-8 shows that the average time required by a student to finish his daily assignment varied from 50% to nearly 100% of the time available, with an average of 78%. As will be discussed later, this time correlates very highly with the achievement measures ($r < -.90$) so that the students who did best needed the least time. The average student "saved" more than 20% of his time and the best student about 50% of it. These times were adequate for most students to achieve about the same skill as their non-CAI counterparts. One might attempt to achieve an improved distribution of skill, with the weaker students being more similar to the stronger by changing the work assignments so the weakest student might, e.g., work three or four times as long as the strongest rather than only twice as long as in the present experiment. With such changed work assignments, we would expect to observe the narrower spread of scores between weaker and stronger students that was looked for.

Course Quality

The quality of the experimental course was good enough to demonstrate the usefulness of CAI, and was certainly very good considering that it was the first effort of its type by the curriculum team involved.

Nevertheless, there were two respects in which the course could be substantially improved. One of these (remarked on by authors, teachers, and students alike) is to develop improved exercise forms for use with the advanced material taken up in the latter part of the course. Such developments were not possible in the present project because of contraints of time and resources.

The other is to provide a larger amount of supplementary materials for the weaker students. That was not done in this project because we could not provide extra time in which the weaker students could do the supplementary work. We would estimate that about 30% to 50% more material should be available.

Despite the above remarks the IBM CAI group is of the opinion that the overall quality of the course was quite good, and did not unfavorably affect the outcome of the experiment to any marked extent.

Test Instruments

To have most precisely measured the effectiveness of learning we should in principle have prepared special tests for the CAI students. However, development and calibration of suitably reliable tests would have been quite unfeasible in this experiment. Also, for other reasons, it would still have been necessary to test the

CAI students with the same tests as were used for the non-CAI students, the DLI tests for the regular course, and the MLA Foreign Language Cooperative tests.

We have noted the conventional DLI tests created a disadvantage for the CAI students in a few cases, either because it presented an unfamiliar testing situation or because the testing task involved subskills of a non-linguistic nature not present in the CAI training task. Obviously, comparisons of methodological effectiveness are made difficult where the primary skills we want the student to learn are confounded with secondary skills associated to the testing tasks. However, we are not ultimately concerned with questions of relative effectiveness in the abstract; we want to know *which kind of learning task and secondary skills are most appropriate in view of the requirements if the job the DLI graduate must do*?

To establish the effectiveness of CAI in relation to this ultimate job requirement, not merely in relation to some artificial tasks pursued for training purposes, we would need to consider, for example, to what extent dictation *as such* is used in the final job, whether paper and pencil or typewriter secondary skills are most useful, etc. Unfortunately, we do not have a fully adequate specification of these ultimate objectives of DLI training, so we cannot discuss this point.

Internal Scoring Measures

An important object of this study was to determine the usefulness of internal CAI scores as measures of student progress toward DLI objectives. The usefulness of CAI scores as performance indicators is central to the effective use of CAI as a student management tool. The potentials of CAI management are:

(1) to conserve the individual learner's time by limiting practice on any subskill to the amount needed for mastery.

(2) to prevent the loss of weak learners by helping to avoid drops due to failure.

(3) to permit accelerated progress of the exceptionally strong student.

(4) to permit a minimization of extra staff effort and trainee time required for periodic tests of performance skills.

The key to the use of CAI scores to allocate student effort is the reliability with which the computer can measure the quality of a student response. Good reliability is required both in the evaluation of a specific response and in the building up of an overall index of performance in each important skill areas.

In order to assess the quality of CAI performance indices we must examine how well the CAI indices correlate with the various other criterion tests of known reliability. Thus we need to know the reliability of the criterion tests themselves.

Information as to the reliability of the MLA tests may be gotten from the Educational Testing Service (ETS) *Handbook* and the *Booklet of Norms*. These normalization tables show that these tests, *considered as specialized tests of the skills by which they are designated*, have test-retest reliability of about .90. We will be concerned below with the MLA tests *as measures of overall quality of learning in the DLI environment*; the parallel test reliability of the individual MLA tests *when used for this purpose*, is substantially lower.

No documents similar to the ETS *Handbook* and *Booklet of Norms* exists from which we could get reliability information for the DLI tests. However, a rough estimate of these reliabilities can be obtained from the coefficients of intercorrelation of the components of the DLI test among themselves and with the total test itself. The method of making these estimates is explained in Appendix III-J.

The estimated reliabilities we can obtain are of two sorts. One is an *internal estimate* based on the degree of correlation of a composite score to its components or with a corresponding score at a different time. This internal reliability refers to the reliability of the composite as a measure of whatever the components have in common. The second is a reliability as a predictor of the DLI grades and is based on the correlations of the composite score with the DLI grades.

	R of composite (internal measure)	R of composite (ref to DLI grade)
DLI (test)	.87	.89
DLI (class)	.87	.89
DLI (grade)	.94	.94
MLA (total)	.80	>.84
CAI (time)	.92	>.90
CAI (int)	.92	>.75
CAI (1st att)	.95	>.77

Table III-J-2 Reliabilities of composite tests.

Table III-J-2 shows values for these reliabilities as determined from our data. Several points are worthy of note. The DLI test grade and class grade are of a high reliability (.89) similar to that of the ETS standarized MLA tests. The MLA tests, for which each component has a test-retest reliability of about .90, show low stability from 12 to 18 week test periods, and relatively low intercorrelation of component scores. Thus as tests of general language learning the MLA components are of poor reliability and the MLA composite is of moderate reliability, whether by the internal criterion or by the criterion of correlation with the DLI grades.

The CAI composite skills indices all have high internal reliability, which implies that the skills involved in all four kinds of CAI exercises have large common components. The CAI time variable also correlates very highly with the MLA total and with the DLI academic grades, as highly as, or more highly than, the independent grades correlate with one another. The CAI internal and first attempt scores correlate highly with one another but only moderately well with the DLI grades; they are not quite as reliable as the MLA composite test as a predictor of DLI grade.

The "blocking power" or predictive power of the CAI scores is shown in Figures III-J-3 to III-J-6. In these figures the means of the quartile scores are shown for the 12-week and 18-week final grades using ALAT, 3-week test, and the three composite CAI scores as blocking variables. To determine the data for the figures the surviving students were divided into quartiles on the basis of scores computed for that blocking variable and plotted above its marked position on the abscissa.

The "power" of a blocking variable is indicated in these plots by how far apart are the quartile means of scores on that variable. The figures show that all the blocking variables are capable of separating the top quartile from the other three. However, it is clear that the quartiles as defined by the CAI variables, and especially by CAI (time) are much better separated and farther apart than are the quartiles defined by either ALAT or 3-week test.

This capability of the CAI internal variables to correctly separate students into groups that will have similar performance on the DLI tests can be directly translated into power to improve instruction through student management. It is clear that the students in the lowest quartile _if_ correctly identified at an early time, could be given extra work and shaped up to meet higher standards; however, the key to the success of such an approach is how early and how accurately one can classify a student as weak or strong.

In order to look at the power of the CAI time variable as a current measure we computed the weekly averages of this variable and weekly recency averages with weight 2 (some what like a 3-week sliding average). The correlations of these variables among themselves are shown in tables III-G-13 and III-G-14. For one-week values there is a mean week-to-week correlation of .84, which gives an estimate of the intrinsic reliability of the one-week value of the time variable. Of especial interest is the high persistence of this correlation over many weeks.

In table III-G-15 we show the correlation between current values of these variables and the large composite tests, i.e., the 12th week value with the 12 week tests, etc. These correlations are for the most part of the order .90, again indicating the great power of the time variable as a measure of current performance.

Other comparisons examined but not presented here for lack of time, indicate that by the end of the first week of CAI work, the CAI time variable constituted a very powerful blocking variable for assigning people to quartiles on the 18-week test, *even though during this first week the students were learning to type*.

There is little doubt that further experimentation with the CAI scoring system using indices that combine speed and accuracy measures would result in creation of test variables excellent in every way for control of student progress through the school.

Student Reaction

The students were asked to give a reaction to various technical features of the CAI system and the CAI course by marking a questionnaire and then adding their suggestions as to how the course might be improved. A copy of the questionnaire, a summary of the student responses, and a transcript of their suggestions are included as Appendix H. We will here summarize what we regard as the most significant features of student reaction.

The students found the equipment acceptable on the whole, but found the exercises occasionally confusing. They thought the exercises more helpful than not by a ratio of 7 to 1 with one neutral; at the same time they found the exercises neither especially interesting or especially dull for the most part. They thought CAI work distinctly more effective than homework, and were about equally divided as to whether it was most useful in the beginning of the course or throughout. They were fairly positive in their overall assessment of CAI as a language learning aid and on the whole would prefer to be in a CAI section rather than a regular section; however, they (correctly) did not think they would make better grades than the non-CAI students. The suggestions of the students as to how to improve the course were on the whole quite good and confirmed some of the views of the DLI staff who prepared the experimental program. Among the suggestions:

1. More emphasis on aural comprehension.
2. Improved techniques of dealing with aural comprehension.
3. Relate the CAI lab to previous day's work rather than current new assignments.
4. Change the dictation format to make it agree with that of the test situation. (There was much comment on the test situation.)
5. A new daily schedule using some CAI, some regular lab.
6. Deletion of name recognition exercise and translation. (There is some evidence that the reaction against translation is especially

directed to the last part of the course where problems of ambiguity were serious).

As a whole the student reaction is sensible and credible, since they found CAI worthwhile but not glamorous. Their acceptance where appropriate would seem to be based on its being useful and a worthwhile change of activity from other kinds of instruction.

III. Discussion of Results: operational questions

Before military training language programs can take full advantage of the instructional potential of CAI extended methods development and instructional program preparation must be carried out in a pilot operations environment. One major consideration in beginning such future pilot operations in CAI will be to achieve economically practicable approaches to the problems of instructional programming and operational use. Thus two primary goals of the present project were to permit HQ, DLI to evaluate the difficulties inherent in program preparation and of the operational integration of the CAI units into a multimedia language course.

The operational statistics show that the present project demonstrated a satisfactory mastery of these problems for purposes of pilot operations. The ad hoc programming system used, although crude in some ways, permitted the curriculum team to achieve a high productivity of instructional materials. Making allowances for the unusual skill and motivation of the curriculum team and recognizing that additional effort would be required to take the materials through several cycles of revision beyond what was done in this experiment, it would seem that after some method development CAI instructional materials could be produced by selected

DLI staff at operationally practicable costs, assuming that costs are acceptable if they can reasonably be amortized over a few thousand students.

A second area of practical interest is the ease of integration of the CAI units into the multimedia language course. In the present experimental situation the DLI course manager successfully administered the CAI lab as an instructional component in a similar way as any other instructional component. A single computer professional was able to do all necessary liaison with the machine center and to coordinate various technical phases of the operation. This operation, too, ran as smoothly as is necessary for a pilot operation.

It is true that some further technical-administrative problems would have to be dealt with if one had a full-shift prime-time CAI system that was more highly integrated into the school. However, results of the present experiment indicate that pilot operations could be reasonably undertaken on the basis of state of the art techniques.

Operational Projections

IV. Effort Analysis and Cost Projections

As a result of the operational experience in this project it is possible to make a realistic analysis of the major components of effort that should be important cost factors in a future operational use of CAI by DLI. These cost factors are of two types, the costs of preparing instructional material and the costs of operational instruction.

In order to establish a framework for cost projections we go through two steps of analysis;

(A) we analyze the effort expended in the present experiment into appropriate categories;

(B) we project the corresponding effort in each category for a hypothetical operational environment.

From the effort projections cost projections can be made directly using whatever institutional costing rates are applicable.

A. Effort analysis of experimental program.

Throughout all phases of the experiment records were kept of the effort expended by both IBM and DLI. This effort can be analyzed into three components: 1) the design and implementation of the CAI control program, 2) the preparation and compilation of instructional material, and 3) the actual CAI operations. The second category comprised five subcategories, including

1) management, 2) authoring, 3) clerical, 4) technicians, and 5) machines, with the effort distributed between an initial authoring phase and a revision phase. The magnitude of the individual components of effort expressed in man hours or equipment hours are itemized below:

Category 1: Preparation of Control Program

Programming (professional man hours)	680
Compiling (machine time)	111

Category 2: Preparation of Linguistic Material

	Initial Authoring	Revision
Management--Professional		
Linguistic Authoring	307	27
Technical	280	160
	587	187
Authoring--Professional		
Text	637	147
Tape Recording	47	4
	684	151
Professional Totals	1271	338
Professional hours per hour of instructional materials	10.1	

	Initial Authoring	Revision
Clerical--Non-Professional (text and tape script typing and keypunching)	428	80
Technicians--Non-Professional (tape recording and marking)	92	8
Military Personnel (SP 4th class for audio tape debugging)	172	98
Non-Professional Totals	692	186
Non-Professional hours per hour of instruction	5.5	

	Initial Authoring	Revision
Machines		
CAI System Access	280	160
1050 Keypunch System	280	160
7010 Compiling	286	164

Category 3: CAI Operations (Seventy-one Daily CAI Lessons)

	Hours
Technical Management	568
Proctor	568
Machine and Communications	
10-1050 AV's (to 7010 computer)	4 months
10 off-premise IBM Telephone extensions	4 months

B. Operational Projections

On the basis of the foregoing effort analysis we will project the following major components of effort in a hypothetical CAI operational installation:

1) developing control programs for foreign language CAI
2) authoring foreign language CAI instructional materials and courses
3) operating a CAI instructional system

Preparing the Control Program

Assumptions:

1) Future CAI programs will continue to consist of groups of exercises of a limited number of formal types.
2) Future CAI language programs will be similar to present programs in the level of complexity of the control program and the language processing load on the computer.
3) Future CAI programs will be prepared in some high level language like Coursewriter, COBOL, FORTRAN, or PL/I.
4) Future CAI systems will have terminal capabilities generally similar to those used in this experiment, and will be operated under time sharing systems of high efficiency.

Control Program Preparation:

The size and complexity of the control program for a CAI foreign language course is determined entirely by the number of types of exercises and the complexity of the processing they involve, and not at all by the length of the course as such. Furthermore, a single control program, once perfected, could be used with only minor changes for a number of different courses. Thus in an operational environment the effort involved to prepare control programs is to be measured by how often a major revision is to be made in the method of instruction and the exercise types; it is not directly related to the number of students serviced by the school.

For greatest efficiency of production and quality of the product, the authoring of linguistic materials should be done in the context of a control program of known design using a specialized programming aid, which we will can a lesson compiler. The lesson compiler is revised with the control program.

The control program used in the DLI-IBM experiment, providing for ten exercise types, can serve as a measure of the programming effort required to produce a control program for a typical foreign language CAI program if we assume:

1) a similar number of statements per exercise
2) similar type number of man hours and machine

hours per statement

3) a contingency of 50% for revisions

4) for each three professional hours one clerical hour for key-punching and documentation (not all required in this project)

5) a lesson compiler will be written to accompany the control program.

Production of Instructional Course Materials

In estimating the effort to produce CAI instructional materials we will assume:

1) The CAI program will be related to other course materials including texts.

2) First version materials will be prepared in large units (e.g., one hundred hours of CAI instruction) by an intensive effort of a curriculum team.

3) Within a CAI program exercises will be grouped into blocks of a single type.

We would project somewhat more time and materials per hour of instruction than was expended on the present project; both because we feel these authors achieved productivity above the normal and had inadequate time for making needed revisions of material, and because in a more general CAI situation one would need to prepare somewhat more supplementary materials.

Effort Estimates:

A low and a high estimate are given to indicate the approximate range of costs, depending on both work effectiveness and scope of goals.

	low estimate	high estimate
1) Control program preparation		
Initial program		
programming effort	800 man hrs.	1200 man hrs.
clerical	250 "	375 "
machine time (compiling)	125 hrs.	187hrs.
Annual maintenance	20% of above	
2) Materials of instruction		
per hour of instructional material		
Initial running version		
professional (teacher/mgr.)	8	12
professional (computer)	2.2	3.3
technical/clerical	5.5	8.3
First revision		
professional (teacher/mgr.)	4	8
professional (computer)	1.1	2.2
technical/clerical	2.8	5.6
Annual maintenance	20% of initial cost	

Projected operational costs for a hypothetical CAI Operation at DLI

Near term total costs of operating a complete CAI system are rather large, both because of the quasi-experimental nature of such systems, and because available systems use computers that are of substantial size. Precision in projecting costs for hypothetical future operational CAI instructional systems is limited by the fact that at present there is no computer industry product system that is clearly a prototype of a future standard CAI product system.

In projecting costs of the components of a future commercial CAI system on the basis of components of present computer systems one gets an estimate that is probably too large, since it does not take into account possible cost savings from volume production; this point is especially important in the relation to terminal equipment.

Near Term Cost Productions

To our best belief at least three vendor companies, viz. IBM, Philco-Ford, and Radio Corporation of America, have product systems that are capable of performing some or most of the functions executed by the CAI system used in the current project. We do not have the technical information necessary to compare the relative suitability of these systems; even if we did, accurate cost performance comparisons would require extensive study, since system performance is quite sensitive to details of system architecture. For these reasons our near-term projections will be based just on the IBM systems, for which our knowledge and understanding is best.

Configuration A

Of present IBM systems, system costs for the DLI application are best defined for the direct-access, 1500 time-sharing system, a non-commercial CAI system which IBM had made available on a restricted basis to selected education institutions for experimental and evaluational purposes. A system configuration promising relatively favorable cost/performance (operation cost per terminal hour basis) in the DLI environment might consist of the following:

a) an IBM 1130 central processing unit, Model 3D, or an IBM 1800 central processing unit.

b) 30-IBM 1510 Instructional Displays (cathode ray tube, keyboard, and light pen,

c) 30-IBM 1506 Audio Units

d) 30-IBM 1512 Image Projectors (optional)

e) 2-IBM 1518 Typewriters

f) various necessary system components

In an operating environment, the system will accommodate 30 students at a given time.

A system having lower cost but less favorable cost/performance could be configured with e.g., only 20 units 1510, 1506, and 1512. Further by limiting exercise types it would be possible to dispense with some 1512's.

The staff for such a 1500 CAI computation center would have to provide for machine operation and computer software support. Presupposing a 50-hour instructional week (10 hours per day, 5 days per week) and 28-hours per week of "background work" (4 hours daily Monday-Friday, 8 hours Saturday), an adequate computation center staff might consist of:

a) one computation center director,
   1) to manage center operations,
   2) to assume overall technical responsibility for the production and maintenance of CAI course materials,
b) two machine operators, one working on swing shift,
c) one systems programmer responsible for the maintenance of system programs
d) one CAI programmer
   1) to develop and maintain CAI control programs,
   2) to modify existing control programs, and
   3) to assist and consult in course production

The system could provide, for example, 2 hours daily of CAI instruction to each of 150 students per week. At such a student load two or three language courses could be supported.

Computation Center Permanent Staff

a)	Computation Center Director	1700 hrs/yr
b)	Machine Operators (2)	3400 hrs/yr
c)	Systems Programmer	1700 hrs/yr
d)	CAI Programmer	1700 hrs/yr
	TOTAL	8500 hrs/yr

CAI Operations

Proctors (2)	2500 hrs

System

IBM 1500 Computer System with 30 student stations and fast CPU including one time charges	200,000/yr
IBM 1500 Computer System with 20 student stations	130,000/yr

Configuration B

An alternate system approach to setting up a CAI center might be for the CAI users to time share a large computer with other users. Such an arrangement would be attractive if CAI methods were to be integrated with other uses of data processing equipment in connection with administrative and instructional functions of management; it may well be the predominant mode of access to computer facilities at some future time.

Since no integrated product system of processor, terminal equipment, and programs fully equivalent to the IBM 1500 is currently available, we can present only rough estimates of the costs for such a system. We estimate that without including any special engineering or systems programming, a current (third generation) time sharing system with an optimized operating system and terminals similar to those used in the present experiment might have a total system cost of $800-$1,000 per terminal per month. Of this total cost about $100-$150 per month would be communications costs involved in connecting the terminals to the central processor, and the remainder would be about equally divided between terminal and processor costs. These cost assumptions do not take account of the possibility that a user may have a central computer that is not completely utilized;

in such a case his true net cost for the CAI might actually be set lower, perhaps as low as $400-500 per month per terminal.

Long Term Cost Projections

Since there does not exist at this time an established commercial standard for a CAI system, projections of long term CAI costs are highly speculative. They involve estimates of the cost/performance of future processors and of future terminals. Future terminal costs will depend sensitively on the volume of manufacturing of appropriate terminals which will itself depend on how rapidly the use of such terminals grows.

We do not wish to predict when large volume commercial CAI will be available or what will be its ultimate level of costs. An indication of the possibilities for systems costs has been given by the CAI pioneers, Drs. Alpert and Bitzer of the University of Illinois, who have projected for the 4,00 terminal CAI system which the Illinois PLATO project will have operational in the mid '70's a system cost of the order of $.30 per terminal hour.

A 4,000-terminal system dedicated to CAI would clearly be unsuitable for most users. Nevertheless, if it were feasible for smaller users to share such a system and realize the same low price/performance figure, the system cost for a CAI installation would be reduced by about a factor ten over the costs cited in our short term estimates.

It may be too optimistic to assume such low costs for smaller users by the mid '70's. It might be better to think in terms of a more price/performance improvement for smaller-than-university-scale CAI, perhaps to achieve system costs of the order of $1.00 to $1.50 per terminal hour on commercial time shared CAI systems in the mid '70's.

At future times when equipment costs are markedly lower than at present, the structure of the total system cost will change drastically. As equipment costs become a more modest fraction of the total, the system cost will become sensitive to the ratio instructional load divided by course production load. Also, while early projects will necessarily be burdened with the cost of producing instructional materials, production costs should become a smaller part of the whole as the school accumulated a library of acceptable programs. Moreover, when a larger system gets to be justified the ratio of staff costs to machine costs could be smaller, since the number of highly skilled computer professionals need not be increased with the number of terminals.

These various factors are combined to produce the following estimates of system costs for two sizes of installation, which would serve for e.g., one course to each of about 500-750 and 1500-2150 DLI students, respectively. These installations would all have somewhat similar costs for computer technical staff.

Total Equipment Costs Estimates

One year system cost	low estimate	high estimate
20-30 terminal		
1970-73	$130,000	$200,000
1975-80 (small user)	50,000	115,000
1975-80 (PLATO)	15,000	22,000
90 terminal		
1975-80 (small user)	$200,000	$350,000
1975-80 (PLATO)	66,000	66,000

Future Directions

V. Directions for Future Work

The work of the present project establishes that within at least a part of the DLI environment CAI techniques can be effective aids to language learning and reliable tools for evaluating student skill in basic language mechanics. It also shows that the present implementation of CAI could serve as a prototype of a pilot CAI system that would be well adapted to DLI faculty needs.

Several possible future benefits from the use of CAI can be postulated on the basis of this work. These benefits would come essentially from the improved utilization of staff time and student time that computer aids would make possible. Possible sources of time savings include acceleration of learning through individualized student management, reduction of time used for testing, and reduction of teacher-trainee contact time.

The magnitudes of any such gains are purely speculative at the present time. Moreover, given the cost of available systems it is doubtful that the large scale introduction of computer methods would result in immediate substantial cost savings. However, the ultimate potential is real and sufficient to justify DLI in undertaking a CAI development program at an appropriate time to prepare to take advantage of future cost-benefits.

Recommendation for a Pilot Operational Installation

Even without complete cost-justification on basis of its instructional value alone, a pilot CAI installation for operational research and development would be well justified at the present time as a vehicle for investigating the values that could be obtained from the use of CAI in the DLI environment. *It is recommended that such a pilot operation be established as soon as operational priorities permit.*

Detailed suggestions for future work

Future CAI development work in DLI should stress all of the following themes:

1) Improved techniques of interactive CAI fully adapted to the goals of military language training.
2) Exploitation of the built-in testing power to accelerate better students, increase student survival rates.
3) Integration of CAI performance information into the management information system of the school.

As an aid to developing a future CAI installation, we have included in this report a body of information about the exercises, control programs, and production techniques that worked out successfully in this experiment.

We will now discuss some detailed ideas about where future work should be done differently, where more should be done, and where, although nothing was done in this study, something should be done.

Exercise Types

1) Several of the exercise types used in this study such as dictation should be altered so that time pressure is present. The preferred approach would be to treat each first attempt as a test, and allow student self-pacing only on subsequent tutorial attempts. Another approach might be to have alternate forms of the exercise, some of which are administered under time pressure.

2) Most of the present exercises are best suited for use in the early parts of the course; they need to be modified to be fully suitable for the advanced material. In addition the invention of new exercise types for advanced students is needed. To give one example: toward the end of the course simple translation exercises become seriously ambiguous because of the great number of possible answers; at the same point in the course exercises involving some translation might be more suitable for grammar exercises than the simple substitution exercises used earlier. One might do well to design an advanced grammar exercise combining features of the simple translation and grammar exercises.

3) A new class of exercises should be studied which would make use of visual displays as part of the cuing system.

4) Some study should be given to improved formulas for scoring the exercises. A simple scoring system based on percentage of correct characters was used in the experiment, and worked acceptably. However, the quality of scoring of grammar and translation exercises could be sharpened up by providing the teacher a means to indicate that only certain *portions of a sentence* are to be examined for correctness; such an improvement is easily possible for all the exercises within the present CAI approach. In addition improvements of the quality of scoring could be made by incorporating timing measures related to those involved in the CAI time variable into some of the scoring formulas.

Lesson Design

1) We have several times remarked on the need for changed lesson design as the students move from basic mechanics to the more advanced parts of the course. This change should be reflected not merely in improved exercise types, but also in improved exercise mix. In the latter part of the course the student needs not only both more complex training tasks, e.g., that require larger retention span, more highly structured

behavior, etc., but also a different distribution of student time among the types of training tasks.

2) The instructional environment should permit extra work sessions and more variation to individual assignments. When these advances can be made, different algorithms should be used in the control programs, although the general approach of having obligatory and supplementary modules is basically a good one.

Course Organization and Course Management

The greatest opportunities for getting additional instructional power from CAI require changes in the course organization and the management of the classes. When students can be sorted into fast and slow sections, it will clearly be best for all if each of these sections goes through the course at such a pace as to optimize its rate of progress while maintaining a high standard of proficiency.

The present course organization, to the extent that it is based on a fixed daily quantum of presentation and review, is not well adapted to an administrative arrangement in which sections progress at varying rates according to their ability. However, the present DLI course materials could be adapted to such a variable rate system (with some changes in course structure) if one recognizes that it already contains identifiable units of about ten days' duration that are the natural units for such a system. If the administration of these present larger units were changed to concentrate presentation of new material in the beginning and to concentrate review at the end, it would be possible by varying the number of review units from section to section to adjust to the different rates of progress of sections of different ability.

The development and test of such a variable progress course organization is recommended.

Testing and Test Development

1) In implementing a variable speed system the use of daily on-line practice in the CAI mode could play the crucial role of generating the test data that is used to pace the section. The present study shows the feasibility of an approach in which CAI testing would make the large periodic DLI tests largely unnecessary, so they could be either largely done away with, or perhaps replaced by a more specialized test of a possibly new type.

2) A program of test adaptation and/or new test development should be carried on as part of the development of a variable speed system.

3) If course management procedures allow varying periods of transit through the course, a number of

further changes in the management of the school would become appropriate. Training cost assessment and early prediction of training costs could be done on an individual basis since students would no longer have the same training period, or the same training costs. More frequent transfer of students between sections would be feasible, since with sections moving at different rates there would be a more nearly continuous distribution of classes; thus when a student had to drop back for one reason or another, he need not in general lose six weeks as at present, but perhaps only two or three. Management could obtain and make use of a weekly status report on each student based on CAI data, stating, in effect, just how he would stand if he had just completed a large test such as is currently given at six week intervals. Finally the capability to make continual statistical analyses of student and group performance would be of great value to the entire DLI operation as a means to supervise, control, and improve the school.

The development of a management information system component making use of student performance data is recommended.

Course Content Control

A major need of the DLI school is a means to summarize, categorize, and evaluate the specific content of its courses. It is clear that the use of data processing systems in DLI would permit the relatively rapid preparation of thesauri and concordances to be used to characterize existing materials and as an aid to revision. Exploitation of such text processing methods would create the potential for frequent systematic reviews and revisions of content, a potential that is recognized to be very important, but is presently not feasible.

As part of the process of introducing computer aids to instruction it would by most desirable to plan how to introduce such basic text processing methods into the curriculum production and maintenance process. Thus a set of programs and procedures should be developed for analyzing and displaying information about the linguistic materials in existing courses, for preparing tables of linguistic materials as aids to authors in construction and revision of course material, and for checking and summarizing linguistic data in an author's current work as an aid to efficient production of instructional material.

The development of a suitable text processing system for use in evaluating, revising and preparing courses of all types is recommended.

TABLE OF CONTENTS OF THE APPENDICES

Page

Appendix I--Phase I 1

- Sec. A Design of Experimental Course 3
- Sec. B Exercise Types 33
- Sec. C Segments of Production Manual 45
- Sec. D Text and Tape Scripts 55

Appendix II--Phase II 75

- Sec. A Subject Selection 76
- Sec. B Aptitude Comparisons 84

Appendix III--Phase II 89

- Sec. A Experimental Design 91
- Sec. B Tables of Grades and Scores 101
- Sec. C 12-week Performance Comparisons 111
- Sec. D 18-week Performance Comparisons 125
- Sec. E Results of Principal Analyses 139
- Sec. F Tabulation of Means 151
- Sec. G Correlations and Intercorrelations 159
- Sec. H Student Questionnaire 175
- Sec. J Test Reliability, Formulae, Tables 191
- Sec. K 24-week Performance Comparisons 201

Appendix I

APPENDIX I

Phase One

Section A

This section describes the design of the experimental AC course.

Section A

PHASE I-COURSE DESIGN AND PRODUCTION

THE AURAL COMPREHENSION COURSE

The Russian AC course was a thirty-seven week course of instruction attended by military personnel six hours each day in classes of nine or ten students. The major unit of instruction is the *daily cycle* of six hours. In the conventional AC course, these six hours are allocated to the following activities:

1st hour--Instruction of new grammar with reinforcing structural drill

2nd hour--Dialogue presentation with reinforcing activities

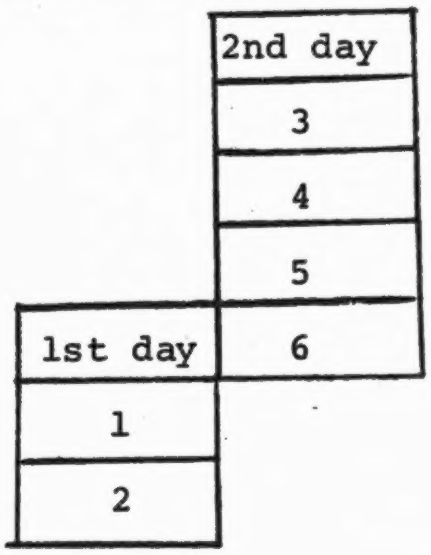
3rd hour--Dialogue recitation

4th hour--Grammar drill based on first hour material and new lexical content

5th hour--Reading/narration and English to Russian translation

6th hour--Dictation and aural comprehension

The cycle is spread across two days with the last four hours of the cycle occurring on the second day, as illustrated in the diagram below.



The rationale for this distribution is based on the view that independent study of the material presented in the first two hours of the cycle is desirable if the student is to gain the maximum amount from the remaining four hours. Generally, the independent study (homework) requires memorization of a dialogue and writing out a number of grammar and vocabulary oriented exercises. Except on days devoted to review prior to tests, a new lesson with new grammatical and lexical material is presented every day. For four hours of the cycle, students meet with an instructor; two hours are spent in a conventional audio language laboratory. In the AC, one of those lab hours is devoted to the presentation of the dialogue associated with the lesson; the other hour involves the student in various exercises in aural comprehension, e.g., several varieties of dictation, sentence comprehension, discourse comprehension and others.

The AC course is organized into 6-week blocks with a major examination being administered at the end of each of these periods. In addition, an examination is given at the end of three weeks, by which time students have been introduced to the phonology and Cyrillic orthography. The 3-week test, the grades for which are not recorded, is used by ESD primarily to indicate those students who appear to be having extreme difficulty. The 6-week tests have several components. These are (1) dictation, (2) oral production, (3) reading, and (4) aural comprehension, and (5) grammar, selected in various combinations for particular 6-week tests.

A number of grades are recorded for each student in order to facilitate close supervision of student progress. First, a daily class grade is recorded which includes homework and class performance. Secondly, the averaged scores for the six week test are recorded. Third, an overall class grade computed from daily class grades, 6-week tests, and previous course grades.

Students are selected for AC on the basis of interest (all are volunteers) and the results of the Army Language Aptitude Test.

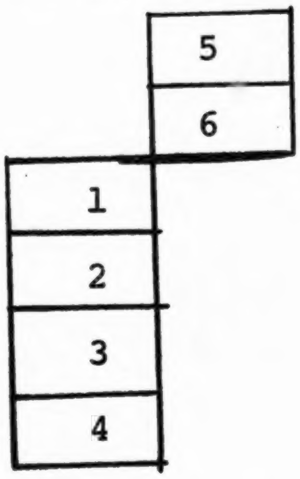
SUBSTITUTION OF CAI FOR UNITS OF THE DAILY CYCLE

The CAI system designed at IBM Research for foreign language instruction was intended to provide supervised practice in the production and comprehension of foreign language sentences. The DLI AC curriculum provides for interactive practice between student and teacher in a number of different forms. Some of these appeared to be more suitable than others for CAI implementation given the present state of CAI techniques. The segments of the daily cycle in which the most extensive use of suitable exercises was made proved to be the Grammar Drill and Dictation and Aural Comprehension hours. Thus, the chosen experimental arrangement was one that involved replacement of these hours by two CAI lab hours.

Two factors necessitated the rescheduling and restructuring of the DLI daily cycle for that group of students selected for CAI participation. Owing to the facts (1) that the CAI computer was located in Yorktown Heights, New York and (2) telecommunications cost had to be minimized, it was necessary to make use of the transcontinental IBM dial network for data transmission during the period after the end of the business day in the East, that is, after 2PM Pacific Time. This meant that CAI equipment would be available to students during hours largely outside of the daily cycle, which ends at 4PM.

The second factor was a time efficiency factor. The two CAI hours could be most effectively used if they ran consecutively because such an arrangement would minimize administrative overhead time, e.g., signing students on the system, setting up tapes, signing off, etc. This consideration suggested the need to restructure the cycle by making the grammar drill and special aural comprehension hours contiguous.

The restructured daily cycle ultimately adapted for the CAI group was as follows:



1st hour--Introduction of new grammar with reinforcing structural drill (class)

2nd hour--Dialogue presentation with reinforcing activities (class)

3rd hour--Grammar drill (CAI-Audio Lab)

4th hour--Dictation and aural comprehension (CAI-Audio Lab)

5th hour--Dialogue recitation (class)

6th hour--Reading/narration and English to Russian translation (class)

The operational schedule for the CAI group was to be as follows:

		M	T	W	Th	F
Class	8-9	DR 8	DR 9	DR 10	DR 11	rev. or test
	9-10	R&N 8	R&N 9	R&N 10	R&N 11	rev. or test
	10-11	GP 9	GP 10	GP 11	rev.	GP 12
	1-2	DP 9	DP 10	DP 11	rev.	DP 12
CAI	2-3	GD 9(I)	GD 10(I)	GD 11(I)	ACC rev.(I)	GD 12(I)
	3-4	ACC 9(I)	ACC 10(I)	ACC 11(I)	ACC rev.(I)	GD 12(II)
	4-5	GD 9(II)	GD 10(II)	GD 11(II)	ACC rev.(II)	GD 12(III)
	5-6	ACC 9(II)	ACC 10(II)	ACC 11(II)	ACC rev.(II)	
	6-7	GD 9(III)	GD 10(III)	GD 11(III)	ACC rev.(III)	
	7-8	ACC 9(III)	ACC 10(III)	ACC 11(III)	ACC rev.(III)	

In addition, for special reasons, the following variation from the daily schedule were planned.

1. Friday afternoon would have an abbreviated schedule of one hour. This hour would typically be devoted to grammar drill on the new lesson.

2. Thursday afternoons which start review cycles would have extra AC practice and grammar review.
3. Special review days would have appropriate special combinations of exercises.
4. Thursday of the first week of the operational phase (third week of the course) would have a special late one-hour orientation lesson.

CONTENT OF THE CAI PROGRAM

The curriculum of the AC course is composed of daily units which are either lessons in which new material is presented, or reviews of lessons. In all, seventy-one daily units had to be created for the experimental course. The content of this CAI component of the course was adapted directly from the existing AC texts since the intent of the entire experiment was to vary methodology with a minimum change of content. Thus, the CAI exercises were designed to provide suitable practice in the lexical and grammatical features in the correct sequence.

Exercise Types

Language instructors have devised numberous language learning tasks for use in face to face exercises. Some, but not all of these, are potentially appropriate for the CAI laboratory. The exercise types used in IBM programs satisfied two criteria thought to be especially important: 1) the learning task should require the student to take a creative action in the target language, e.g., either to generate an utterance or to construct a response based on comprehension of an utterance, and 2) the learning task should evoke behaviors that employ the critical components of the subject matter.

Through properly designed CAI exercises a language learner should develop both a knowledge of how the sentences of the language are constructed, what the words mean, and how the meaning of sentences is extracted from sentence structures, and the ability to construct and comprehend the sentences of the language in speaking, listening, reading, and writing.

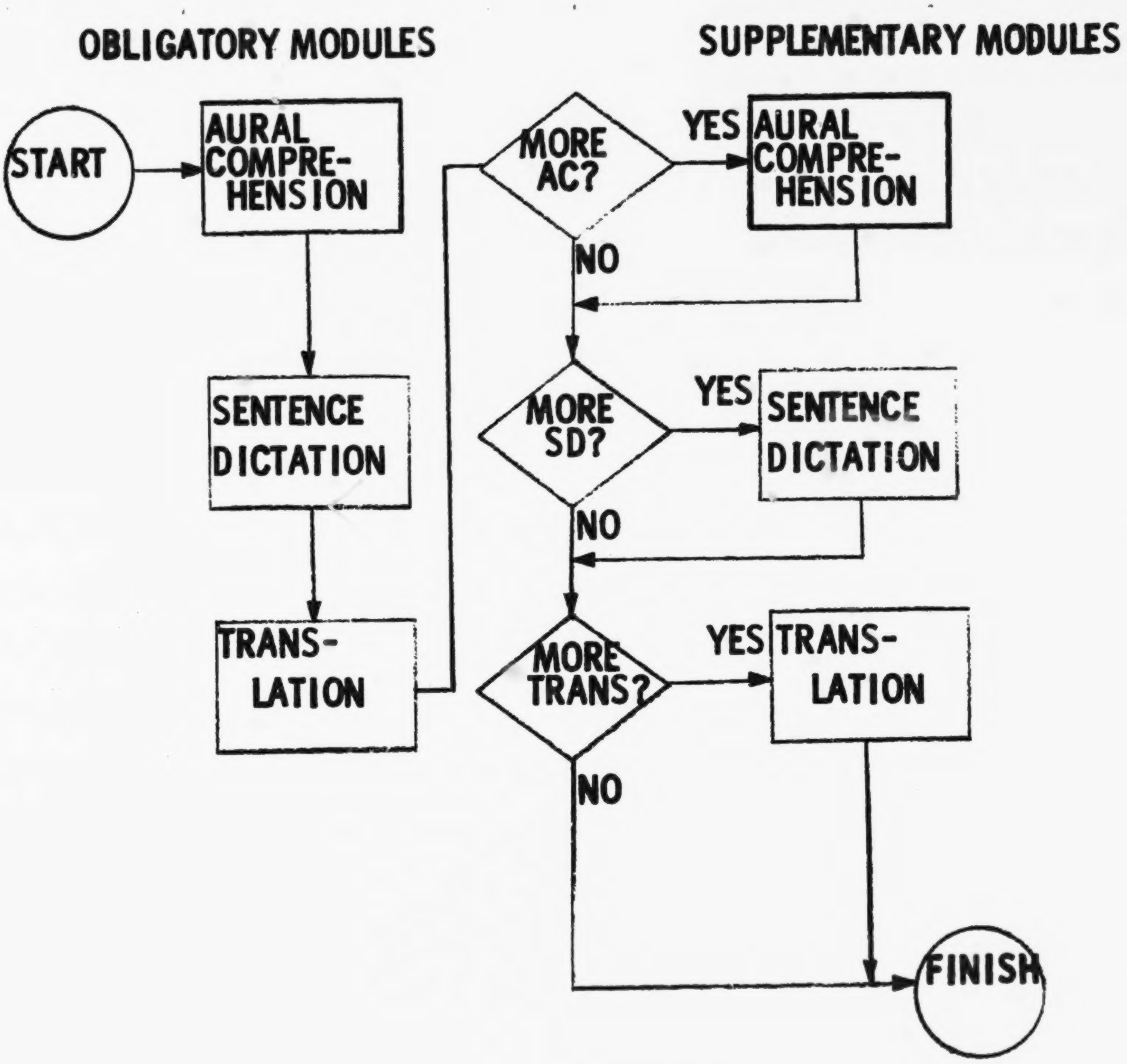
Ten exercise types were selected as best suited to accommodate both the subject matter materials and the instructional methods. These exercises are described together with their rationale in Appendix I, Section B.

Lesson Design

In a particular lesson, a number of different CAI exercises occur. Exercises of the same type are grouped into sets called modules, each module containing four or five items. A lesson consists of an obligatory block of modules and a supplementary block. Every module in an obligatory block is paired with one of the same type in a supplementary block, a translation with a translation, a dictation with a dictation, and so forth. Supplementary modules were intended to be used only by students who needed extra work in that particular mode, this decision being made by the computer on the basis of how well the obligatory work was performed.

This course of events in an instructional situation is illustrated roughly in the diagram below. In this hypothetical abbreviated lesson, the student first takes a module of aural comprehension exercises, then a module of sentence dictation exercises, finally a module of translation exercises. At this point he has completed the obligatory modules.

Now control passes to the supplementary part of the program. The machine tests the student's score on aural comprehension; if it meets criterion, he will not be required to do the supplementary aural comprehension module; if it is below criterion, he will be assigned the supplementary aural comprehension. The same pattern is followed for the sentence dictation and translation supplementary modules. When the student passes through all of the supplemental work required on the basis of his scores, he has finished the lesson.


OBLIGATORY MODULES
SUPPLEMENTARY MODULES
START
AURAL COMPREHENSION
SENTENCE DICTATION
TRANSLATION
MORE AC?
YES
NO
AURAL COMPREHENSION
MORE SD?
YES
NO
SENTENCE DICTATION
MORE TRANS?
YES
NO
TRANSLATION
FINISH

In all, five different daily module arrangements were used. These are described in Appendix I, Section B, Part I.

PEDAGOGICAL CONTROL PROGRAM

The supervisory component of the CAI program may be called the pedagogical control program, or, simply the pedagogy. This program administers the scoring rules and makes four different kinds of decisions regarding the flow of control. Two of these decisions relate to repetition of an attempt on an item without scoring, a third relates to item termination, the last relates to possible assignment of supplementary modules. The nature of the control decisions will now be described, and then the scoring rules and procedures on which they are based.

Attempt Pedagogy

1) Defaulting an Attempt--Under various circumstances, it is desirable to default a student response, permitting him to retry the task without having been scored on his previous attempt. Such circumstances arise typically when a student miskeys or accidentally inputs irrelevant material to the system. This can happen either when a student is careless or in rare instances where a faulty audio tape will cause a message (e.g., a dictation sentence) to be played to the student which does not correspond to what the system expects as the correct answer.

When defaulting a student answer, the system transmits the simple message "repeat," leaving the student to decide whether (a) he has been careless but is reprieved and allowed without penalty to enter his message again or, (b) the audio tape is at fault, so he should initiate the tape adjustment procedure.

The computer recognizes the need to default a student input when the score on the first attempt at any item is "too low," that is, below a teacher prescribed criterion value. In this case, the input is scrubbed and a "repeat" message is sent. The criterion value in the present study was maintained at a 25% degree of match between the student response and the correct answer.

2) Tape Replays--Allowing the student replays of audio tape messages is a desirable feature of audio oriented CAI programs and this feature was incorporated into the DLI program. To hear a tape message again, the student merely keys a certain symbol on the keyboard and the tape message is replayed. In order to encourage students to perfect comprehension skills, it was felt to be desirable to be able to limit the number of tape replays available to a student. Consequently, a count on tape replays within an item is kept on each attempt and

compared, whenever updated, with a criterion value. If ever that value is exceeded, the student is informed that no further requests for tape replays will be honored.

Item Pedagogy

After each attempt at an item, it must be decided whether the student should make another attempt or whether the particular item should be terminated. In the DLI CAI program, item termination comes about under a variety of different circumstances.

1) Number of Attempts--Students are usually allowed more than one attempt on an item in the DLI program. However, the number of attempts is limited so as to avoid continued practice under conditions of diminishing return. In the DLI program, the maximum number of attempts allowed to the student is three. If, on the third attempt, the student has not gotten the answer correct, the program supplies the correct answer to him and passes on to the next item.

2) Item Correct--When the student gets any attempt correct, his success is confirmed for him and the item is terminated.

3) High Proficiency--If a student's performance is very good, it is deemed unlikely that he will profit much from repeating that item. Thus when the student's attempt score exceeds a given "escape value," e.g., 95, the system terminates the item by giving the student the fully correct answer and passing him on to the next item.

4) Skip Option--The student has the option to terminate any item rather than make a new attempt, regardless of his score. This option is thought valuable for a number of reasons. First, there may be errors obvious or subtle in the program, so that it is impossible or unnatural for the student to give the answer expected. Or the student may simply not see the point, so that he can't work on it. In either case it is inefficient or irritating if the student is forced to enter a message that will be treated as a bona fide attempt at the item. Thus, the student is permitted to skip any item.

When a student skips an item his score on the item is fixed at the value already generated by his work on it. Thus if he exercises the skip option on his first attempt, he is given a zero for the item.

Module Pedagogy

With the completion of an obligatory module, a program decision is made as to whether the student will be required to take the corresponding supplementary module. This decision is based upon two scores, the module score and the lesson recency score. If either of these scores falls below an acceptable minimum, a switch is set in the computer memory, with the consequence that the student will subsequently be required to take the supplementary module.

Program Parameters

The entire DLI course is structured around a set of daily lesson programs, each of which corresponds to a particular daily unit of instruction. There are seventy-one such unit programs. In addition, there is a master control program which directs the student to the appropriate lesson. Finally, there is a very small "setup" program which establishes the values of parameters used in each daily unit. Parameter values are passed to the appropriate daily unit program just at the beginning of execution of that program.

Parameters are established in an independent program rather than in the daily units so as to minimize the compiling time needed to make changes in the parameter values controlling some or all of the daily units. Compile time for the parameters program is approximately six minutes. Compile time for a single daily unit program, on the other hand, averages seventy-five minutes.

The control parameters set by the setup program are:

The Course Weight--The value of the weighting factor employed in the computation of the course recency score. Initial value = 9.

The Lesson Weight--The value of the weighting factor employed in the computation of the lesson score. Initial value = 2.

The Item Weights--The value of the weighting factors employed for the first, second, and third attempts in the computation of the item score. Initial values = 3-3-4 respectively.

The Default Score--This is the value which is compared with the student's attempt score to determine whether the attempt should be defaulted and the student be requested to try again. Initial value = 25.

High Proficiency Score--This is the value which is compared with the student's attempt to determine whether he did sufficiently well to be allowed to skip on to the next item. Initial value = 97.

Remediation Escape Score--This value is compared with the module score and the lesson recency score to determine the necessity for the student to do the corresponding supplementary module. Initial value = 95.

Maximum Number of Attempts--This value specifies the number of attempts allowed to a student on a given item. Initial value = 3.

Maximum Number of Tape Replays--This value specifies the number of tape replays allowed the student on a particular item. Initial value = 5.

Exercise Group Scoring

Exercise types fall into four general skill categories. These are 1) grammar exercises, 2) translation, 3) dictation exercises, and 4) aural comprehension.

1) Grammar Drill Exercises:
    - Textual Substitution-Transformation
    - Aurally Cued Substitution-Transformation
    - Aurally Cued Transformation

2) Translation (English to Russian)

3) Dictation:
   - Name Dictation
   - Number Dictation
   - Specialized Vocabulary Dictation
   - Sentence Dictation

4) Aural Comprehension:
   - Aural Comprehension of Sentences
   - Aural Comprehension of Discourses

In order to allow evaluation of the relative contributions of the four exercise groups to overall student performance, the CAI program provides a separate scoring system for each of the exercise types, thus four distinct sets of program parameters and four distinct sets of cumulative scores, e.g., four module scores, four lesson recency scores, four course recency scores, etc., as described below.

The Scoring System

The scores generated in the CAI program are designed to serve two main purposes: one, to provide a basis for machine decisions about the flow of control; the other, to provide indices of student performance for use by human course supervisors. In its mechanics the scoring system follows somewhat different plans at different levels, resulting in five kinds of scores corresponding to different types of control decisions or performance indices, *viz.*, attempt scores, item scores, module scores, lesson scores, and course scores.

Attempt Scores

An *attempt* is defined as a single student response on a particular task and each student attempt is given a score. The attempt score is derived from two values. The first is the percentage of match between the characters in a student response and the characters in the correct answer. This value is generated by the special partial answer function which constructs partial answer feedback. The second value is an excess character percentage which is non-null whenever the partial answer value exceeds 100%.

Schematically, the attempt score is derived as follows:

$$\text{Score}_{\text{attempt}} = \begin{matrix}\text{partial}\\ \text{answer}\\ \text{percentage}\end{matrix} - \frac{1}{2}\ \frac{\begin{matrix}\text{no. of characters in}\\ \text{excess of correct ans.}\end{matrix}}{\text{characters in correct ans.}}$$

It is seen that this particular equation weights excess character errors one half as heavily as missing characters in assessing degree of correctness.

Item Scores

An item score is a weighted average of attempt scores based upon at most the first three attempts at the item. Each attempt score contributes a percentage of the item score in proportion to the weights assigned in the scoring rule. The scoring parameters are readily adjusted and are normally changed from time to time in operational circumstances for experimental and other reasons. The general working of the weighted average item scoring system is given by the following rule:

$$\text{Score}_i = \frac{\left(\begin{matrix}\text{1st} & & \text{1st}\\ \text{att} & \times & \text{att}\\ \text{sc} & & \text{wgt}\end{matrix}\right) + \left(\begin{matrix}\text{2nd} & & \text{2nd}\\ \text{att} & \times & \text{att}\\ \text{sc} & & \text{wgt}\end{matrix}\right) + \left(\begin{matrix}\text{3rd} & & \text{3rd}\\ \text{att} & \times & \text{att}\\ \text{sc} & & \text{wgt}\end{matrix}\right)}{10}$$

An example may make clear the working of this system. Suppose that the following weights are used:

1st attempt weight = 3
2nd attempt weight = 3
3rd attempt weight = 4

The student's score on the first attempt thus fixes 30% of his item score; the second attempt score another 30%; and the third attempt score, the remaining 40%. If, for one reason or another, either the student or the computer terminates the exercise before the third attempt (e.g., the student gets the problem completely right on the first attempt), the score for all non-completed attempts is set to be the same as the score of the most recently completed attempt, as is seen in the generalized item score formula below.

The item score S_i is computed when the ith item is completed after A attempts.

$$S_i = \frac{1}{10} \sum_{a=1}^{A=1} s_a w_a + s_A \sum_{a=A}^{\infty} w_a$$

The w_a are positive integers, their total is 10. In our program only w_1, w_2, and w_3 are non-zero. The w_a tell how the different scores are weighted when the student makes several attempts on a single item.

By suitably varying the w_a a teacher can adjust the amount of credit given for improvement on the second and third attempts. To illustrate further the use of the w_a, consider the following additional examples. Let $w_1 = 2$, $w_2 = 5$, and $w_3 = 3$.

Example 1. On one item a student makes four attempts, attaining scores $s_1 = .60$, $s_2 = .70$, $s_3 = .90$, and $s_4 = .95$. His score is

$$S = \frac{2(.60) + 5(.70) + 3(.90)}{10} = .74$$

Note that the w_a values 2, 5, and 3 signify that 20% of the score was fixed on the first attempt, 50% on the second, and the remaining 30% on the third attempt. The performance on the fourth attempt did not influence the final item score.

Example 2. If the w values had been the same as above but the student had stopped after two attempts, his final item score would be

$$S = \frac{2(.60) + (5 + 3)(.70)}{10} = .68$$

In this case, 20% of his score was fixed on the first attempt, and the remaining 80% of his score was fixed on the second attempt because it was the last.

Module Score

After the full complement of items in a module are completed, a module score is computed. In the present experimental program the module score is the simple arithmetic average of the items scores for items contained in the module.

$$Sc = \frac{item_1 + item_2 + \ldots + item_n}{N}$$

Lesson Recency Score

The lesson recency score is a cumulative score intended to reflect how well the student has been doing on modules of a particular type. This score is computed at the completion of each module by the following recursive procedure:

$$\text{New Score}_{les} = \frac{(\text{Old Score}_{les} \times \text{recency wgt}) + \text{module score}}{\text{recency wgt} + 1}$$

The recency weight is an integer between 0 and 9 which tells how heavily the old lesson score is to be weighted in combining it with the most recent module score. The greater the value of this weight the greater will be the effect of the old lesson recency score on the new lesson recency score and the slower will be the change in the lesson score with time.

Course Recency Score

The course recency score is computed daily at the end of a CAI daily unit. It is computed recursively in accordance with the following kind of formula:

$$\text{New Score}_{\text{course}} = \frac{(\text{Old Score}_{\text{course}} \times \text{recency wgt}) + \text{lesson score}}{\text{recency wgt} + 1}$$

The course recency score reflects how well students have done over a period of days, the number of days being a function of the particular recency weight employed.

COURSE PRODUCTION

At the outset of the first phase of the project, IBM and DLI worked jointly to design a course production plan providing for integration of the activities of the DLI course authors, the IBM on-site technical manager, and the computation center in Yorktown Heights, New York. This production plan provided, in effect, an author's manual with detailed written specifications for:

1) the exercise types
2) the design of each daily lesson
3) a production schedule
4) a description of authoring procedures
5) a description of interface procedures for communication between authors and the on-site technical manager.
6) an administrative plan for the interfacing of effort betwee IBM and DLI

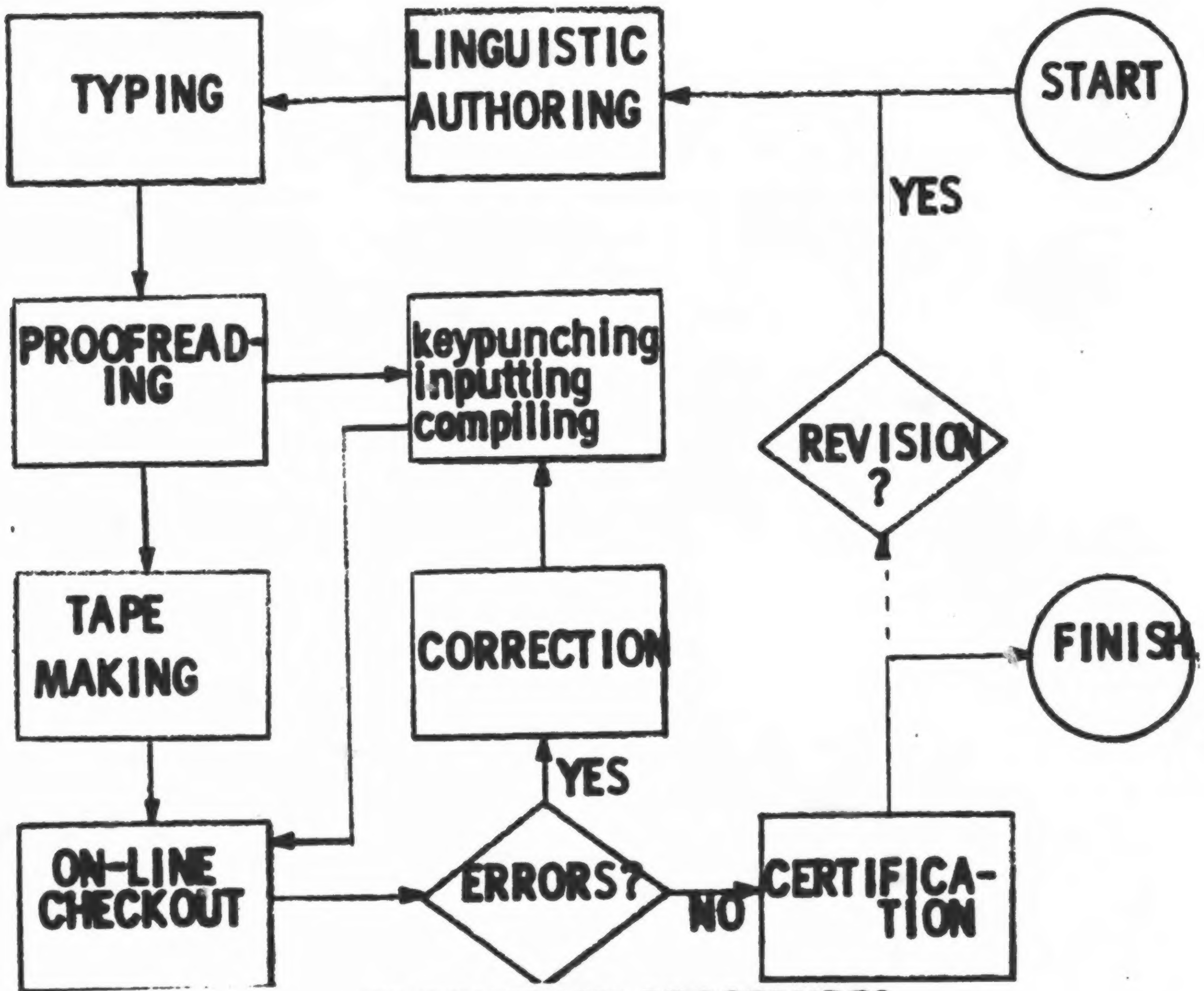
The technical components of the production plan are attached as Appendix I, Section C. The essentials of the production procedures could be summarized as follows:

1) An author prepares a manuscript that contains linguistic materials.
2) The manuscript is typed and proofread and tape scripts are prepared from it.

3) Tapes are made and proofed in parallel, the exercise texts are keypunched, input to the machine, and compiled.

4) The completed program and tapes are checked out operationally. If there are program errors, corrections are input to the machine, and the program given another operational check out.

5) When there are finally no errors, the program is certified as ready to use.


TYPING
LINGUISTIC AUTHORING
START
YES
PROOFREADING
keypunching inputting compiling
REVISION ?
TAPE MAKING
CORRECTION
FINISH
YES
ON-LINE CHECKOUT
ERRORS?
NO
CERTIFICATION

PRODUCTION PROCEDURES

The production schedule required that the first versions of lesson manuscripts all be completed by the middle of the production period, so that there would be time for revision where the DLI staff felt it was most needed. The original schedule was adhered to quite closely, so that revisions were essentially complete and the programs ready to run two days before the end of the production period.

Section B

This section contains a description and linguistic rationale of the exercise types implemented in the CAI program employed in the DLI experiment.

Section B

Description of Exercise Types

1. Textual Substitution Transformation--TST

In a textual substitution transformation exercise the computer types out to the student a "kernel" sentence and, below it, one or more cues.

(computer) я рад, что вы сейчас свободны.

(computer) мы она

The student's task is to type the sentence resulting from the substitution of the cue items into the appropriate positions of the kernel sentence. For example, if the student did this particular exercise correctly, he would type:

(student) мы рады, что она сейчас свободна. п

(computer) --↑

Each TST module (i.e., a collection of items) contained five items and one model item. In the model, the computer would type the kernel, the cues, and the correct answer so that the student might see what was expected of him.

The following special considerations were observed by course authors:

a) Wherever possible, each item contained two cues, as in the example above.

b) Each substitution should normally have required a double transformation of the kernel. For example,

(computer) я рад, что вы сейчас свободны.
(computer) он   она
is an abnormal item because the substitution will result only in the transformation of "свободны."

c) The pattern of substitution should be exactly the same throughout the module. The same constituents should be substituted consistently.

Textual Substitution--Transformation (TST)

(computer) я рад, что вы сейчас свободны.
(computer) мы   она

---

(student) мы рады, что она сейчас свободна. п (Correct)
(computer) --↑

---

(student) мы ради, что она сейчас свободны. (Incorrect)
(computer) мы рад-, что она сейчас свободн-.
(student) мы рады, что она сейчас свободна. п (Correct)
(computer) --↑

2. Aurally Cued Substitution Transformation--ACST

The aurally cued substitution transformation module is identical to the TST module with the exception that the cues are played on the tape

3

recorder rather than be typed by the computer. Aside from this, the special considerations for TST hold for ACST.

Aurally Cued Substitution Transformation (ACST)

(computer) брату пять лет.
(computer) audio message: я двадцать четыре

(student) мне двадцать четыре года. п (Correct)
(computer) --↑

(student) мне двадцать четыре лет. (Incorrect)
(computer) мне двадцать четыре ----.
(student) мне двадцать четыре года. п

3. Aurally Cued Transformation--ACT

In an ACT module, as in the substitution transformation modules, the student is provided with a model and five items. Here, a model consists of a taped kernel after which the computer types out a desired transformation of the kernel, as in:

(computer) audio message: там есть письма для солдат.
(computer) там нет писем для солдат.

In each of the following five items, a tape kernel will be played and the student must type out the desired transformed sentence.

The following special considerations were observed:

a) The sentence structural pattern given in the model item should be constant throughout the entire module. Otherwise, the student does not know what transformation to produce.

b) Wherever possible a transformation should result in multiple changes of the kernel sentence.

c) Sentences should be kept fairly short.

Aurally Cued Transformation (ACT)

```
(computer)  audio message:  там есть стулья для друзей.
-------------------------------------------------------
(student)   там нет стульев для друзей.         п (Correct)
(computer)  --------------------------------------↑
-------------------------------------------------------
(student)   там нет стулов для друзей.           (Incorrect)
(computer)  там нет стул--в для друзей.
(student)   там нет стульев для друзей.         п
(computer)  --------------------------------------↑
```

4. Translation--TRANS

In a translation exercise the computer types a sentence in English to the student, who is then required to type a Russian translation of the sentence. If his translation is correct, the computer will type "п" for "правильно." Otherwise, the computer will type back a partial answer feedback line and the student will reenter his answer. Each translation module consists of five sentences (items) to be translated.

The following special considerations were observed:

a) Sentences should be fairly long and should contain as many elements of the grammar and vocabulary of the particular lesson as is possible.

b) English sentences should be free of ambiguity.

Translation--English to Russian (TRANS)

```
(computer)  my parents live in a village near the airport.
-----------------------------------------------------
(student)   родители живут в селе около аэродрома.  п (Correct)
(computer)  --------------------------------------↑
-------------------------------------------------------
(student)   родители живят в селе около аэродром.  (Incorrect)
(computer)  родители жив-т в селе около аэродром-.
(student)   родители живут в селе около аэродрома.  п
(computer)  ------- ------------------------------↑
```

Translation is a non-paradigmatic exercise, unlike the exercise types mentioned previously. Thus, if a particular lesson should contain a number of small isolated grammatical points which cannot be practiced economically in a paradigmatic fashion, then a translation sentence or two can be employed to address the topic. Also, translation allows vocabulary practice.

These exercises address the domain of linguistic competence efficiently and comprehensively; they also influence the performance skills of reading, writing, and listening, since these are the media in which the student practices. In doing these exercises, therefore, it is to be expected that students will acquire performance abilities as well as linguistic competence.

Among the exercises appropriate primarily for performance skills are dictation and aural comprehension.

5. Sentence Dictation--DICTS

In a sentence dictation exercise, a sentence in Russian is played to the student on the tape recorder, and he is required to type it. Dictation, especially sentence dictation, 1) provides good practice in aural discrimination including stress and intonation, 2) exercises and tests orthography, and 3) strengthens the student's perception of the correspondences between the spoken and written form. The computer response to his answer is the same as for other exercises. In each dictation module there are five items.

The following special considerations were observed:

a) Sentences could be fairly long and stated either in the declarative, interrogative, or imperative mode.

7. Number Dictation--DICTNU

In an item of this type, the student hears a set of numbers with a pause in between, during which time he is to type the numbers correctly.

```
Dictation--Numbers (DICTNU)
(computer)   audio message:   23   45   92   76   14
-------------------------------------------------------
(student)    23   45   92   76   14                   π
(computer)   --------------------------------------↑
-------------------------------------------------------
(student)    23   62   82   76   14                   (Incorrect)
(computer)   23   --   -2   76   14
          Audio Message:   23   45   92   76   14
(student)    23   45   92   76   14                   π
(computer)   --------------------------------------↑
```

8. Aural Comprehension of a Discourse--ACD

Aural comprehension exercises are designed to train a student in the ability not only to perceive but to understand since in order to perform the task required, he <u>must</u> understand and must listen more and more intently until he does so. In the ACD exercise, a student listens to a discourse on the tape recorder. This discourse may be either a dialogue or a narration. At the end of the discourse, the student will take four exercise items. Each item consists of a question about the discourse which the student is to answer with short answers (i.e., non-sentences). On a particular item, if the student's answer is wrong, he will be given a remedial message, e.g., the pertinent sentence from the discourse, and will then be asked the question again. An ACD module consists, then, of a discourse message, four question messages, four remedial messages, and right answers for four items. The questions are presented to the student in advance on a sheet of paper so that he will know what he is to listen for.

The following special considerations were followed:

a) Discourses were approximately twenty seconds long.

b) Remedial messages contained the correct answer in them.

c) It was desirable for the author to make the correct answer as short as possible. The question in the discourse below has several possible answers:

братья павловы работают на заводе.
они работают на заводе.
на заводе.

на заводе is the best answer because it is the simplest and the one most students are likely to write. However, all three are acceptable.

d) As a general rule, questions were constructed so that there were as few correct answers as possible. The maximum number of alternate answers allowed was four.

Aural Comprehension--Discourse(ACD)

(computer) audio message: братья павловы, коля и ваня, работают на заводе. в свободное время они обычно ездят в деревню. сегодня воскресенье, завод не работает, и они рано утром идут на вокзал. до вокзала недалеко, и поэтому они идут пешком. на вокзале сидят люди и ждут поезда. по перрону ходит начальник станции. скоро приходит поезд.

(computer) audio message: где работают братья павловы?

(student) братья павловы работают на заводе. п (Correct)

(computer) -------------------------------------↑

(student) братья павловы работают на станций. (Incorrect)

(computer) братья павловы работают на ------.

(computer) audio message: братья павловы работают на заводе. где работают братья павловы?

(student) братья павловы работают на заводе. п

(computer) -------------------------------------↑

9. Aural Comprehension of Sentences--ACS

A Russian sentence is played to the student followed by a question in Russian based upon the preceding sentence. The student's task is to answer that question. This exercise type is modeled on the "rapid fire drill" of the conventional Russian AC course. Each module of ACS contains five items.

The following special considerations were to be observed:

a) As in the ACD exercise, correct answers were as short and as few in number as possible.

Aural Comprehension--Sentences(ACS)

(computer) audio message: лекции начинаются вечером.
(computer) audio message: когда начинаются лекции?

(student) вечером п (Correct)
(computer) -------------------------------------↑

(student) don't know (Incorrect)
(computer) audio replay
(student) вечером п
(computer) -------------------------------------↑

Dictation--Sentences (DICTS)

```
(computer)  audio message:  я обычно езжу на поезде.
-----------------------------------------------------------
(student)   я обычно езжу на поезде.                    п
(computer)  ----------------------------------------↑
-----------------------------------------------------------
(student)   я обично ежу на поезде.                  (Incorrect)
(computer)  я об-чно е-жу на поезде.
(student)   я обычно езжу на поезде.                    п
(computer)  ----------------------------------------↑
```

6.   Name Dictation--DICTNA

In a name dictation item, the student hears two or three names with a pause in between. His task is to type the names he hears. In a name dictation module, there were five items each consisting of two or three names.

Dictation--Names(DICTNA)

```
(computer)  audio message:  петров гагарин андреевский
------------------------------------------------------------
(student)   петров гагарин андреевский                  п
(computer)  ----------------------------------------↑
-----------------------------------------------------------
(student)   питров гагарин андреевски              (Incorrect)
(computer)  п-тров гагарин андреевски-
          Audio message:  петров гагарин андреевский
(student)   петров гагарин андреевский                  п
(computer)  ----------------------------------------↑
```

Section C

The key technical segments of the CAI materials of instruction production manual used by DLI course authors are contained in this section. The section includes:

1) the unit type inventory, which states the alternative module organizations available for use in Daily CAI lessons.
2) the lesson design which specifies for each day of CAI operations the unit type to be employed and a syllabus reference for selection of AC and grammar materials. Syllabus references refer to chapters of the DLI Aural Comprehension course texts.
3) instructions for entering linguistic materials on IBM-provided text and tape scripts.

Section C

Technical Sections of the Course Production Plan

Part I Unit types

During the CAI portion of the AC course, five different daily schedules of CAI modules were called for. These included nonreview days (Unit Types A, D, and E), review days (Unit Type B), and Fridays (Unit Type C). Given below are the module distributions for each of these unit types.

Type A -- nonreview day Obligatory Modules	TST ACT DICTS TRANS DICTNA ACD ACST
Remedial and Supplementary Modules	DICTS TST TRANS DICTNA ACD ACT ACST
Type B -- review day Obligatory Modules	TRANS TST ACT DICTS TRANS DICTNA ACST
Remedial and Supplementary Modules	TRANS DICTS TST TRANS DICTNA ACT ACST

Type C -- Fridays Obligatory Modules	ACT TRANS ACST
Remedial and Supplementary Modules	TRANS ACT ACST
Type D -- nonreview day Obligatory Modules	TST ACT DICTS TRANS DICTNA ACS ACST
Remedial and Suplementary Modules	DICTS TST TRANS DICTNA ACS ACT ACST
Type E -- nonreview day Obligatory Modules	TST ACT DICTS TRANS IRN DICTNA ACST
Remedial and Supplementary Modules	DICTS TST IRN TRANS DICTNA ACT ACST

Part II Lesson Design

The number in the Grammar and ACC categories are syllabus references which inform the course author as to the location of the linguistic material upon which the CAI lessons are to be based.

UNIT	1	2	3	4	5
DAY	Tu	W	Th	F	M
DATE	12/3	12/4	12/5	12/6	12/9
GRAMMAR	7	8	9	10	11
ACC	-	7	8	-	10
TYPE	C	A	A	C	A

UNIT	6	7	8	9	10
DAY	Tu	W	Th	F	M
DATE	12/10	12/11	12/12	12/13	12/16
GRAMMAR	12	13	R(8-13)	14	15
ACC	11	12	R(8-13)	-	14
TYPE	A	A	B	C	A

UNIT	11	12	13	-	14
DAY	Tu	W	Th	F	M
DATE	12/17	12/18	12/19	12/20	1/6
GRAMMAR	16	R(14-16)	R(7-16)		17
ACC	15	16	R(7-16)		R(7-16)
TYPE	A	B	B		A

UNIT	15	16	17	18	19
DAY	Tu	W	Th	F	M
DATE	1/7	1/8	1/9	1/10	1/13
GRAMMAR	18	R	R	19	20
ACC	17	18	R	-	19
TYPE	A	B	B	C	A

UNIT	20	21	22	23	24
DAY	Tu	W	Th	F	M
DATE	1/14	1/15	1/16	1/17	1/20
GRAMMAR	21	22	R	23	24
ACC	20	21	22	-	23
TYPE	A	A	B	C	A

UNIT	25	26	27	28	29
DAY	Tu	W	Th	F	M
DATE	1/21	1/22	1/23	1/24	1/27
GRAMMAR					
ACC	25	26	R	27	28
TYPE	A	A	B	C	D

UNIT	30	31	32	33	34
DAY	Tu	W	Th	F	M
DATE	1/28	1/29	1/30	1/31	2/3
GRAMMAR	29	30	R	31	32
ACC	28	29	30	-	35
TYPE	E	A	B	C	D

UNIT	35	36	37	38	39
DAY	Tu	W	Th	F	M
DATE	2/4	2/5	2/6	2/7	2/10
GRAMMAR	33	34	R	35	36
ACC	32	33	34	-	35
TYPE	E	A	B	C	D

UNIT	40	41	42	43	44
DAY	Tu	W	Th	F	M
DATE	2/11	2/12	2/13	2/14	2/17
GRAMMAR	R	R	R	37	38
ACC	36	R	R	-	37
TYPE	B	B	B	C	D

UNIT	45	46	47	-	48
DAY	Tu	W	Th	F	M
DATE	2/18	2/19	2/20	2/21	2/24
GRAMMAR	39	40	41	-	42
ACC	38	39	40	-	41
TYPE	E	D	E	-	D

UNIT	49	50	51	52	53
DAY	Tu	W	Th	F	M
DATE	2/25	2/26	2/27	2/28	3/3
GRAMMAR	R	R	R	43	44
ACC	42	R	R	-	43
TYPE	B	B	B	C	D

UNIT	54	55	56	57	58
DAY	Tu	W	Th	F	M
DATE	3/4	3/5	3/6	3/7	3/10
GRAMMAR	45	46	R	47	48
ACC	44	45	46	-	47
TYPE	E	D	B	C	E

UNIT	59	60	61	62	63
DAY	Tu	W	Th	F	M
DATE	3/11	3/12	3/13	3/14	3/17
GRAMMAR	49	50	R	51	52
ACC	48	49	50	-	51
TYPE	D	E	B	C	D

UNIT	64	65	66	67	68
DAY	Tu	W	Th	F	M
DATE	3/18	3/19	3/20	3/21	3/24
GRAMMAR	53	54	R	55	56
ACC	52	53	54	-	55
TYPE	E	D	B	C	E

UNIT	69	70	71
DAY	Tu	W	Th
DATE	3/25	3/26	3/27
GRAMMAR	R	R	R
ACC	R	R	R
TYPE	B	B	B

Part III Authoring Procedures

1. Selection of Text Script Forms

The unit types and lesson plan in sections A and B of this appendix specify the exercises, and the order in which they come for particular daily units. This plan tells the author which Text Script form to use in writing down the linguistic content of each exercise. For example, the term TRANS, in a unit type means that the exercise is a translation exercise, and that the author is to use the Text Script labeled TRANS to write the material for this exercise.

In the lesson plan, the entry for Wednesday 12/4/68, unit II requires the A Type distribution of modules, that is:

TST
ACT
DICTS
TRANS
DICTNA
ACD
ACST

DICTS
TST
TRANS
DICTNA
ACD
ACT
ACST

The authors doing this daily unit will select Text Scripts corresponding to each exercise. they will put them in the order given in the lesson plan, and they will fill out the forms. (Sample Text Scripts are attached at the end of this appendix.)

2. Filling out Text Scripts

The Text Script forms are designed to be used by the typist inputter as well as the course author. They contain information for the course author and also some computer code. However, with but little practice, the author will be able to set down his material on the Text Scripts unimpeded by the sequence numbers and other material on the form relevant to computer coding.

The conventions governing the use of the Text Scripts by the course authors are quite simple.

a. A form must be filled out completely. If an exercise type calls for five items, then five items must be written.
b. An author writes only on the lines provided.
c. What the author writes on a given line is determined by the "key" in the left margin of that line of the Text Script.
d. Key symbols:

m	This particular item is a model for the student. In certain exercises, as TST, ACST, and ACT, the student is provided with a model which tells him the pattern or patterns that he is to employ in the exercise items to follow.
1	This number denotes the first exercise item.
2	This number denotes the second exercise item.
3	This number denotes the third exercise item.
4	This number denotes the fourth exercise item.
5	This number denotes the fifth exercise item.
k	On the line to the right of this symbol, the author writes a kernel sentence, as in TST and ACST.
c	The author writes for TST or ACST.
a	The author writes the correct answer, the one with which the computer will compare the student's answer.
t	What the author writes on this line will be a tape message to the student.
d	The author writes a short discourse to be read to the students in ACD exercises.

r	This is a remedial message in ACD exercises. If the student fails to answer correctly a short answer question based upon the discourse, this remedial message is replayed. The remedial message should be a one sentence excerpt from the discourse containing the correct answer.
q	This is the tape message which asks a student a question based upon the discourse.
s	This is the source sentence in a translation exercise, the sentence in English that the student is to translate into Russian.

3. Characteristics of particular Text Scripts

Translation (TRANS)

On lines with the symbol "s" preceding them, the author will write the English sentence to be translated. On the lines beginning with "a," the author will write the correct Russian answers for these translations. Observe that for each English sentence, space is provided for four acceptable Russian answers. The author is not obliged to provide more than one correct answer for each item and, as a general rule, the fewer the better. However, if alternate answers are required, authors may feel free to use up to four for any one translation item. Observe further that a translation module consists of five items.

Sentence Dictation (DICTS)

On lines labelled "t," the author types the sentence which is to be recorded on tape and played to the student. On lines labelled "a," the author writes the correct answers. In most cases, the "a" following a "t" will be identical to it. The second "a" line may be used in rare cases where alternate spelling forms exist for a word.

Number Dictation (DICTNU)

The same conventions hold here as for sentence dictation exercises, except that the tape message will contain numbers rather than sentences.

Name Dictation (DICTNA)

The same conventions hold here as for sentence dictation exercises, except that the tape message will contain names rather than sentences.

Aural Comprehension of Discourse (ACD)

On the lines following "d" the author writes the discourse which the student will hear on tape. The label "r" denotes the line on which the author should write the remedial message that the student will hear if he fails to answer the question about the discourse correctly on the first time. On the line to the right of the label "q," goes the actual tape question which the student is asked about the discourse. On the following lines labelled "a," the correct answers or single answer to the question are written by the author. Observe that an ACD module contains four items.

Aural Comprehension of Sentences (ACS)

On the line labelled "t" the author writes a sentence and a question based upon that sentence, both of which are ultimately recorded on tape. This exercise is the equivalent of DLI's "rapid fire drill." On the lines labelled "a" the author writes the answers to be accepted by the computer as correct.

Textual Substitution Transformation (TST)

This exercise type contains a model for the student as well as five exercise items. On the line "k-m" the author writes the kernel sentence which is to be typed to the student. On the line "c-m" the author writes the constituents to be substituted into the kernel sentence. Here the author must be careful to line the constituents up underneath the items to be replaced in the kernel sentence. On the line labelled "a-m" the author writes the correct answer resulting from the substitution of the cue items into the kernel. All three of these lines will be typed to the student so that he will know what is expected of him on the exercise items to follow. In the five items, lines labelled "c" are those on which items to be substituted into the preceding correct answer are written. On the line labelled "a" following, the author writes the correct answer for this substitution. Notice that each correct answer becomes the kernel for the following item. Alternate answers should not be required in this type of exercise, but an extra "a" line is provided just in case an alternate correct answer is unavoidable.

Aurally Cued Substitution Transformation (ACST)

This exercise type is virtually identical in format to TST, the only difference being that the cues are not textual but are to be recorded as tape messages.

Aurally Cued Transformation (ACT)

On the line labelled "t-m" the author writes a kernel sentence for a tape message. On "a-m" he writes the correct answer, which will be a sentence resulting from the transformation of the kernel sentence. Both of these will be model or example items in that the tape recorded will play the line "t-m" and the computer will type the line "a-m" thus informing the student what he is to do on subsequent items. On lines labelled "t" in the actual module items, the author writes kernel sentences. On the lines labelled "a" he wries correct answers, i.e., sentences resulting from the transformation of the kernel sentences.

Item Recognition -- Names (IRN)

In the matrix provided on the Text Script the author writes Russian names which he wants the student to recognize. From this matrix, a handout sheet will be made and given to students in the CAI-Audio laboratory. On a line labelled "t" the author writes three of the names, selecting them from the matrix in any order he chooses. On the "a" following, the author writes the numerals corresponding to the correct names in the matrix. There are five items in an IRN module, each item containing three names.

Section D

Text and Tape Scripts

TRANS Text Script Module No. ____________

Author____________________________ Unit No.__________ Date__________

rr -0101-200 tt ld ⓢ_//c9 e
rr -0101-220 tt ty ⓢ ⓢ
s-1 ____ e
rr -0101-240 tt ld ⓢ
a ____ //b1 e
rr -0101-250 tt ld ⓢ
a ____ //b2 e
rr -0101-260 tt ld ⓢ
a ____ //b3 e
rr -0101-270 tt ld ⓢ
a ____ //b4 e

rr -0102-200 tt ld ⓢ_//c9 e
rr -0102-220 tt ty ⓢ ⓢ
s-2 ____ e
rr -0102-240 tt ld ⓢ
a ____ //b1 e
rr -0102-250 tt ld ⓢ
a ____ //b2 e
rr -0102-260 tt ld ⓢ
a ____ //b3 e
rr -0102-270 tt ld ⓢ
a ____ //b4 e

rr -0103-200 tt ld ⓢ_//c9 e
rr -0103-220 tt ty ⓢ ⓢ
s-3 ____ e
rr -0103-240 tt ld ⓢ
a ____ //b1 e
rr -0103-250 tt ld ⓢ
a ____ //b2 e
rr -0103-260 tt ld ⓢ
a ____ //b3 e
rr -0103-270 tt ld ⓢ
a ____ //b4 e

rr -0104-200 tt ld ⓢ_//c9 e
rr -0104-220 tt ty ⓢ ⓢ
s-4 ____ e
rr -0104-240 tt ld ⓢ
a ____ //b1 e
rr -0104-250 tt ld ⓢ
a ____ //b2 e
rr -0104-260 tt ld ⓢ
a ____ //b3 e
rr -0104-270 tt ld ⓢ
a ____ //b4 e

rr -0105-200 tt ld ⓢ_//c9 e
rr -0105-220 tt ty ⓢ ⓢ
s-5 ____ e
rr -0105-240 tt ld ⓢ
a ____ //b1 e
rr -0105-250 tt ld ⓢ
a ____ //b2 e
rr -0105-260 tt ld ⓢ
a ____ //b3 e
rr -0105-270 tt ld ⓢ
a ____ //b4 e

DICTS Text Script Module No. ____________

Author ______________________ Unit No. ____________ Date __________

```
    t t  -0000-110 t t ld ⓢ _//c9 e
    t t  -0101-200 t t ld ⓢ _//c9 e
t-1 ____________________________________________________________
    t t  -0101-240 t t ld ⓢ
a   ____________________________________________________________//b1 e
    t t  -0101-250 t t ld ⓢ
a   ____________________________________________________________//b2 e

    t t  -0102-200 t t ld ⓢ _//c9 e
t-2 ____________________________________________________________
    t t  -0102-240 t t ld ⓢ
a   ____________________________________________________________//b1 e
    t t  -0102-250 t t ld ⓢ
a   ____________________________________________________________//b2 e

    t t  -0103-200 t t ld ⓢ _//c9 e
t-3 ____________________________________________________________
    t t  -0103-240 t t ld ⓢ
a   ____________________________________________________________//b1 e
    t t  -0103-250 t t ld ⓢ
a   ____________________________________________________________//b2 e

    t t  -0104-200 t t ld ⓢ _//c9 e
t-4 ____________________________________________________________
    t t  -0104-240 t t ld ⓢ
a   ____________________________________________________________//b1 e
    t t  -0104-250 t t ld ⓢ
a   ____________________________________________________________//b2 e

    t t  -0105-200 t t ld ⓢ _//c9 e
t-5 ____________________________________________________________
    t t  -0105-240 t t ld ⓢ
a   ____________________________________________________________//b1 e
    t t  -0105-250 t t ld ⓢ
a   ____________________________________________________________//b2 e
```

DICTNU Text Script Module no. ____________

Author ______________________________ Unit No. ___________ Date ___________

t t -0000-110 t t ld ⓢ_//c9 e
t t -0101-200 t t ld ⓢ_//c9 e

t-1 ____
t t -0101-240 t t ld ⓢ

a ____//b1 e
t t -0101-250 t t ld ⓢ

a ____//b2 e

t t -0102-200 t t ld ⓢ_//c9 e

t-2 ____
t t -0102-240 t t ld ⓢ

a ____//b1 e
t t -0102-250 t t ld ⓢ

a ____//b2 e

t t -0103-200 t t ld ⓢ_//c9 e

t-3 ____
t t -0103-240 t t ld ⓢ

a ____//b1 e
t t -0103-250 t t ld ⓢ

a ____//b2 e

t t -0104-200 t t ld ⓢ_//c9 e

t-4 ____
t t -0104-240 t t ld ⓢ

a ____//b1 e
t t -0104-250 t t ld ⓢ

a ____//b2 e

t t -0105-200 t t ld ⓢ_//c9 e

t-5 ____
t t -0105-240 t t ld ⓢ

a ____//b1 e
t t -0105-250 t t ld ⓢ

a ____//b2 e

DICTNA Text Script Module No. ____________

Author ______________________________ Unit No. ______________ Date ____________

[t t] -0000-110 [t t] ld ⓢ _//c9 🄴
[t t] -0101-200 [t t] ld ⓢ _//c9 🄴

t-1 ____
[t t] -0101-240 [t t] ld ⓢ
a ____ //b1 🄴
[t t] -0101-250 [t t] ld ⓢ
a ____ //b2 🄴

[t t] -0102-200 [t t] ld ⓢ _//c9 🄴
t-2 ____
[t t] -0102-240 [t t] ld ⓢ
a ____ //b1 🄴
[t t] -0102-250 [t t] ld ⓢ
a ____ //b2 🄴

[t t] -0103-200 [t t] ld ⓢ _//c9 🄴
t-3 ____
[t t] -0103-240 [t t] ld ⓢ
a ____ //b1 🄴
[t t] -0103-250 [t t] ld ⓢ
a ____ //b2 🄴

[t t] -0104-200 [t t] ld ⓢ _//c9 🄴
t-4 ____
[t t] -0104-240 [t t] ld ⓢ
a ____ //b1 🄴
[t t] -0104-250 [t t] ld ⓢ
a ____ //b2 🄴

[t t] -0105-200 [t t] ld ⓢ _//c9 🄴
t-5 ____
[t t] -0105-240 [t t] ld ⓢ
a ____ //b1 🄴
[t t] -0105-250 [t t] ld ⓢ
a ____ //b2 🄴

ACD **Text Script** **Module no.** ______________

Author ______________________________ **Unit No.** ______________ **Date** ____________

r r -0000-110 t t ld ⓢ_//c9 e

d ____
____
____
____
____
____

r-1 ____

q ____
r r -0101-200 t t ld ⓢ_//c9 e
r r -0101-240 t t ld ⓢ
a ____//b1 e
r r -0101-250 t t ld ⓢ
a ____//b2 e
r r -0101-260 t t ld ⓢ
a ____//b3 e
r r -0101-270 t t ld ⓢ
a ____//b4 e

r-2 ____

q ____
r r -0102-200 t t ld ⓢ_//c9 e
r r -0102-240 t t ld ⓢ
a ____//b1 e
r r -0102-250 t t ld ⓢ
a ____//b2 e
r r -0102-260 t t ld ⓢ
a ____//b3 e
r r -0102-270 t t ld ⓢ
a ____//b4 e

r-3 ____

q ____
r r -0103-200 t t ld ⓢ_//c9 e
r r -0103-240 t t ld ⓢ
a ____//b1 e
r r -0103-250 t t ld ⓢ
a ____//b2 e
r r -0103-260 t t ld ⓢ
a ____//b3 e
r r -0103-270 t t ld ⓢ
a ____//b4 e

r-4 ____

q ____
r r -0104-200 t t ld ⓢ_//c9 e
r r -0104-240 t t ld ⓢ
a ____//b1 e
r r -0104-250 t t ld ⓢ
a ____//b2 e
r r -0104-260 t t ld ⓢ
a ____//b3 e
r r -0104-270 t t ld ⓢ
a ____//b4 e

r-5 ____

q ____
r r -0105-200 t t ld ⓢ_//c9 e
r r -0105-240 t t ld ⓢ
a ____//b1 e
r r -0105-250 t t ld ⓢ
a ____//b2 e
r r -0105-260 t t ld ⓢ
a ____//b3 e
r r -0105-270 t t ld ⓢ
a ____//b4 e

ACS Text Script Module No. ________

Author ______________________________ Unit No. ____________ Date ________

[r r] -0000-110 t t ld ▪ _//c9 [c]

t-1 ____

____

[r r] -0101-200 [t t] ld ▪ _//c9 [c]

[r r] -0101-240 t t ld ●

a ____//b1 [c]

[r r] -0101-250 t t ld ●

a ____//b2 [c]

[r r] -0101-260 t t ld ●

a ____//b3 [c]

[r r] -0101-270 t t ld ●

a ____//b4 [c]

t-2 ____

____

[r r] -0102-200 t t ld ● _//c9 [c]

[r r] -0102-240 t t ld ●

a ____//b1 [c]

[r r] -0102-250 t t ld ●

a ____//b2 [c]

[r r] -0102-260 t t ld ●

a ____//b3 [c]

[r r] -0102-270 t t ld ●

a ____//b4 [c]

t-3 ____

____

[r r] -0103-200 t t ld ● _//c9 [c]

[r r] -0103-240 t t ld ●

a ____//b1 [c]

[r r] -0103-250 t t ld ●

a ____//b2 [c]

[r r] -0103-260 t t ld ●

a ____//b3 [c]

[r r] -0103-270 t t ld ●

a ____//b4 [c]

t-4 ____

____

[r r] -0104-200 t t ld ● _//c9 [c]

[r r] -0104-240 t t ld ●

a ____//b1 [c]

[r r] -0104-250 t t ld ●

a ____//b2 [c]

[r r] -0104-260 t t ld ●

a ____//b3 [c]

[r r] -0104-270 t t ld ●

a ____//b4 [c]

t-5 ____

____

[r r] -0105-200 t t ld ● _//c9 [c]

[r r] -0105-240 t t ld ●

a ____//b1 [c]

[r r] -0105-250 t t ld ●

a ____//b2 [c]

[r r] -0105-260 t t ld ●

a ____//b3 [c]

[r r] -0105-270 t t ld ●

a ____//b4 [c]

TST **Text Script** **Module No.** ____________

Author ______________________ **Unit No.** __________ **Date** _________

$\boxed{\text{r r}}$ -0100-510 $\boxed{\text{t t}}$ ty ⓢ ⓢ $\boxed{\text{p}}$ a

m-k $\boxed{\text{p}}$ b,, $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0100-520 $\boxed{\text{t t}}$ ty ⓢ ⓢ

m-c ,, $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0100-530 $\boxed{\text{t t}}$ ty ⓢ ⓢ $\boxed{\text{p}}$ a

m-a $\boxed{\text{p}}$ b $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0101-200 $\boxed{\text{t t}}$ ld ⓢ_//c9 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0101-220 $\boxed{\text{t t}}$ ty ⓢ ,, ⓢ

c-1 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0101-240 $\boxed{\text{t t}}$ ld ⓢ

a //b1 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0101-250 $\boxed{\text{t t}}$ ld ⓢ

a //b2 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0102-200 $\boxed{\text{t t}}$ ld ⓢ_//c9 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0102-220 $\boxed{\text{t t}}$ ty ⓢ ,, ⓢ

c-2 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0102-240 $\boxed{\text{t t}}$ ld ⓢ

a //b1 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0102-250 $\boxed{\text{t t}}$ ld ⓢ

a //b2 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0103-200 $\boxed{\text{t t}}$ ld ⓢ_//c9 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0103-220 $\boxed{\text{t t}}$ ty ⓢ ,, ⓢ

c-3 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0103-240 $\boxed{\text{t t}}$ ld ⓢ

a //b1 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0103-250 $\boxed{\text{t t}}$ ld ⓢ

a //b2 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0104-200 $\boxed{\text{t t}}$ ld ⓢ_//c9 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0104-220 $\boxed{\text{t t}}$ ty ⓢ ,, ⓢ

c-4 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0104-240 $\boxed{\text{t t}}$ ld ⓢ

a //b1 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0104-250 $\boxed{\text{t t}}$ ld ⓢ

a //b2 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0105-200 $\boxed{\text{t t}}$ ld ⓢ_//c9 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0105-220 $\boxed{\text{t t}}$ ty ⓢ ,, ⓢ

c-5 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0105-240 $\boxed{\text{t t}}$ ld ⓢ

a //b1 $\boxed{\text{e}}$

$\boxed{\text{r r}}$ -0105-250 $\boxed{\text{t t}}$ ld ⓢ

a //b2 $\boxed{\text{e}}$

ACST Text Script Module No. ________

Author ______________________ Unit No. __________ Date __________

```
      r r   -0000-110  t t  ld  ⓢ //c9  e
      r r   -0000-520  t t  ty  ⓢ ⓢ p  a
k-m   ____________________________________________ p b,,  e
c     ____________________________________________
      r r   -0000-630  t t  ty  ⓢ ⓢ p  a
a     ____________________________________________ p b,,  e

c-1   ____________________________________________
      r r   -0101-200  t t  ld  ⓢ //c9  e
      r r   -0101-240  t t  ld  ⓢ
a     ____________________________________________ //b1  e
      r r   -0101-250  t t  ld  ⓢ
a     ____________________________________________ //b2  e

c-2   ____________________________________________
      r r   -0102-200  t t  ld  ⓢ //c9  e
      r r   -0102-240  t t  ld  ⓢ
a     ____________________________________________ //b1  e
      r r   -0102-250  t t  ld  ⓢ
a     ____________________________________________ //b2  e

c-3   ____________________________________________
      r r   -0103-200  t t  ld  ⓢ //c9  e
      r r   -0103-240  t t  ld  ⓢ
a     ____________________________________________ //b1  e
      r r   -0103-250  t t  ld  ⓢ
a     ____________________________________________ //b2  e

c-4   ____________________________________________
      r r   -0104-200  t t  ld  ⓢ //c9  e
      r r   -0104-240  t t  ld  ⓢ
a     ____________________________________________ //b1  e
      r r   -0104-250  t t  ld  ⓢ
a     ____________________________________________ //b2  e

c-5   ____________________________________________
      r r   -0105-200  t t  ld  ⓢ //c9  e
      r r   -0105-240  t t  ld  ⓢ
a     ____________________________________________ //b1  e
      r r   -0105-250  t t  ld  ⓢ
a     ____________________________________________ //b2  e
```

ACT Text Script

Module no. ______________

Author ______________________________ Unit No. __________ Date _________

r r -0000-110 t t ld ⓢ _//c9 e

t-m ____

r r -0000-620 t t ty ⓢ ⓢ

a-m ____ e

t-1 ____

r r -0101-240 t t ld ⓢ

a ____ //b1 e

r r -0101-250 t t ld ⓢ

a ____ //b2 e

t-2 ____

r r -0102-240 t t ld ⓢ

a ____ //b1 e

r r -0102-250 t t ld ⓢ

a ____ //b2 e

t-3 ____

r r -0103-240 t t ld ⓢ

a ____ //b1 e

r r -0103-250 t t ld ⓢ

a ____ //b2 e

t-4 ____

r r -0104-240 t t ld ⓢ

a ____ //b1 e

r r -0104-250 t t ld ⓢ

a ____ //b2 e

t-5 ____

r r -0105-240 t t ld ⓢ

a ____ //b1 e

r r -0105-250 t t ld ⓢ

a ____ //b2 e

IRN Text Script

Module no. __________

Author____________________ Unit No. __________ Date__________

11	16	21
12	17	22
13	18	23
14	19	24
15	20	25

[t t] -0000-110 [t t] ld ⓢ _//c9 [e]

t-1 ______________________________

[r r] -0101-240 [t t] ld ⓢ

a ______________________________ //bl [e]

t-2 ______________________________

[r r] -0102-240 [t t] ld ⓢ

a ______________________________ //bl [e]

t-3 ______________________________

[r r] -0103-240 [t t] ld ⓢ

a ______________________________ //bl [e]

t-4 ______________________________

[r r] -0104-240 [t t] ld ⓢ

a ______________________________ //bl [e]

t-5 ______________________________

[r r] -0105-240 [t t] ld ⓢ

a ______________________________ //bl_[e]

DICTS Tape Script

AUTHOR____________________________ DATE____________

Unit_______ Module______

Pos.	Tape Message
	This is tape message # . Type plus two.
	Type plus one.
	EOB to begin.
	Type minus one.
	Type minus two.

DICTNU Tape Script

AUTHOR________________________ DATE____________

Unit__________ Module__________

Pos.	Tape Message
	This is tape message # . Type plus two.
	Type plus one.
	EOB to begin.
	Type minus one.
	Type minus two.

DICTNA Tape Script

AUTHOR______________________________ DATE______________

Unit________ Module________

Pos.	Tape Message
	This is tape message # . Type plus two.
	Type plus one.
	EOB to begin.
	Type minus one.
	Type minus two.

ACD Tape Script

AUTHOR____________________________________DATE___________

Unit____________ Module__________

Pos.	Tape Message
	This is tape message # . Type plus two.
	Type plus one.
	EOB to begin.
	Type minus one.
	Type minus two.

ACS Tape Script

AUTHOR________________________________DATE__________

Unit____________ Module________

Pos.	Tape Message
	This is tape message # . Type plus two.
	Type plus one.
	EOB to begin.
	Type minus one.
	Type minus two.

ACST Tape Script

AUTHOR______________________ DATE__________

Unit________ Module________

Pos.	Tape Message
	This is tape Message # ____ . Type plus one.
	EOB to begin.
	Type minus one.
	Type minus two.

ACT Tape Script

AUTHOR______________________ DATE__________

Unit________ Module________

Pos.	Tape Message
	This is tape message # . Type plus one.
	EOB to begin.
	Type minus one.
	Type minus two.

IRN Tape Script

AUTHOR________________________________ DATE____________

Unit__________ Module___________

Pos.	Tape Message
	This is tape message # . Type plus two.
	Type plus one.
	EOB to begin.
	Type minus one.
	Type minus two.

Appendix II

APPENDIX II

Phase Two

Section A

This section contains a graphic account of

1) how the students were grouped and selected for the experimental and control groups
2) how the two groups compared on the predictor variables of ALAT and 3-week test.

Fig. II-A-1 shows all ACC students plotted by ALAT and 3-week test.

Fig. II-A-2 shows the initial grouping of ACC students into thirty groups.

Fig. II-A-3 shows the initial selection of CAI students.

Fig. II-A-4 is a comparison plot of the initial two sections on ALAT and 3-week test. Plotted are rank ordered raw scores on the two predictor measures with variable weighting for non-CAI students. In a group, containing one CAI and one non-CAI, the CAI has a weight of one; in a group of one CAI and two non-CAI students, each non-CAI is weighted one-half.

Fig. II-A-5 shows the results of the regrouping made after the six-week test and necessitated by academic drops from the CAI class at the time of this test.

Fig. II-A-6 is a comparison plot of the regrouped sections on ALAT and 3-week tests.

Fig. II-A-7 shows the group status at the completion of the operational phase.

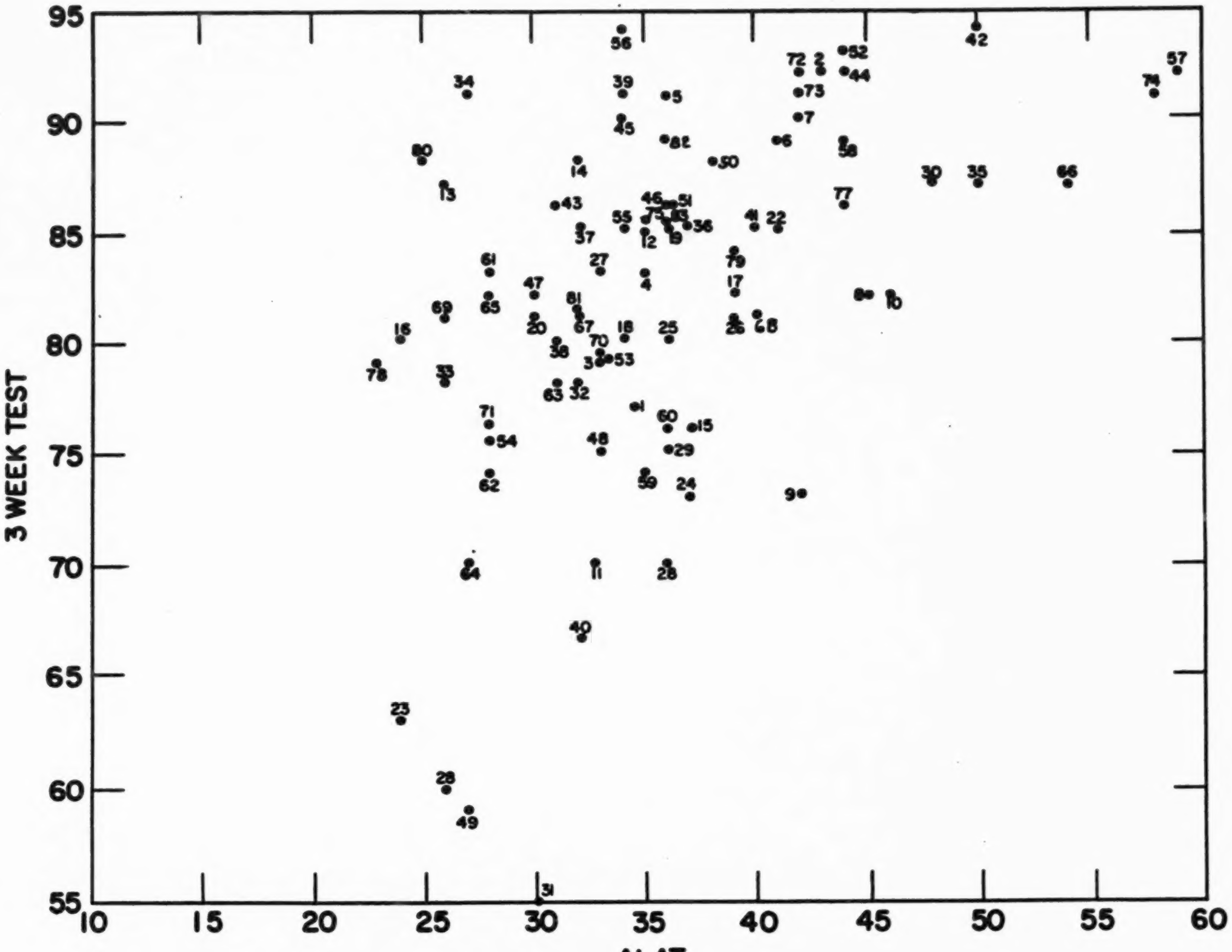
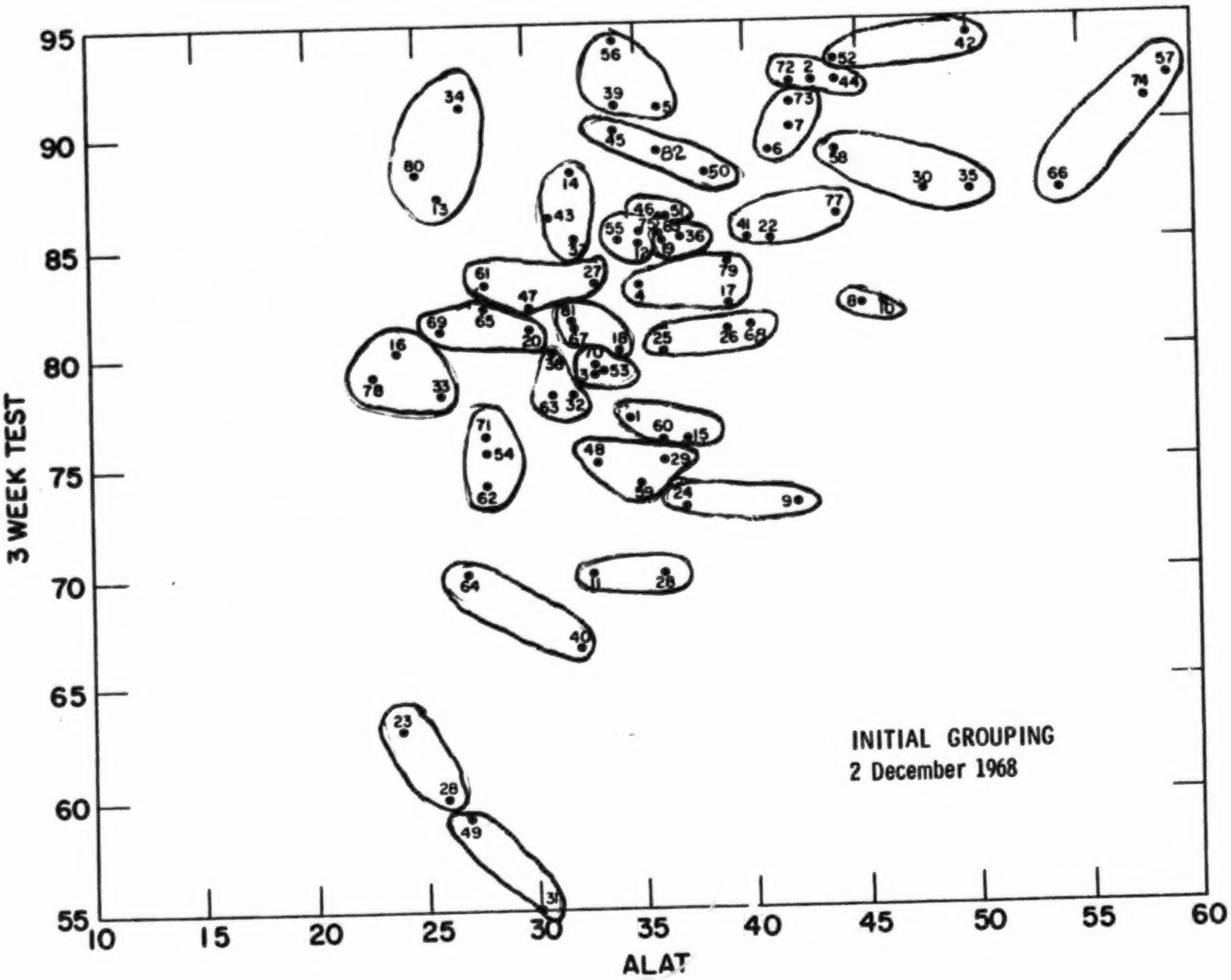


Figure II-A-1 - 3-Week Test vs. ALAT
2 December 1968



INITIAL GROUPING
2 December 1968

Figure II-A-2

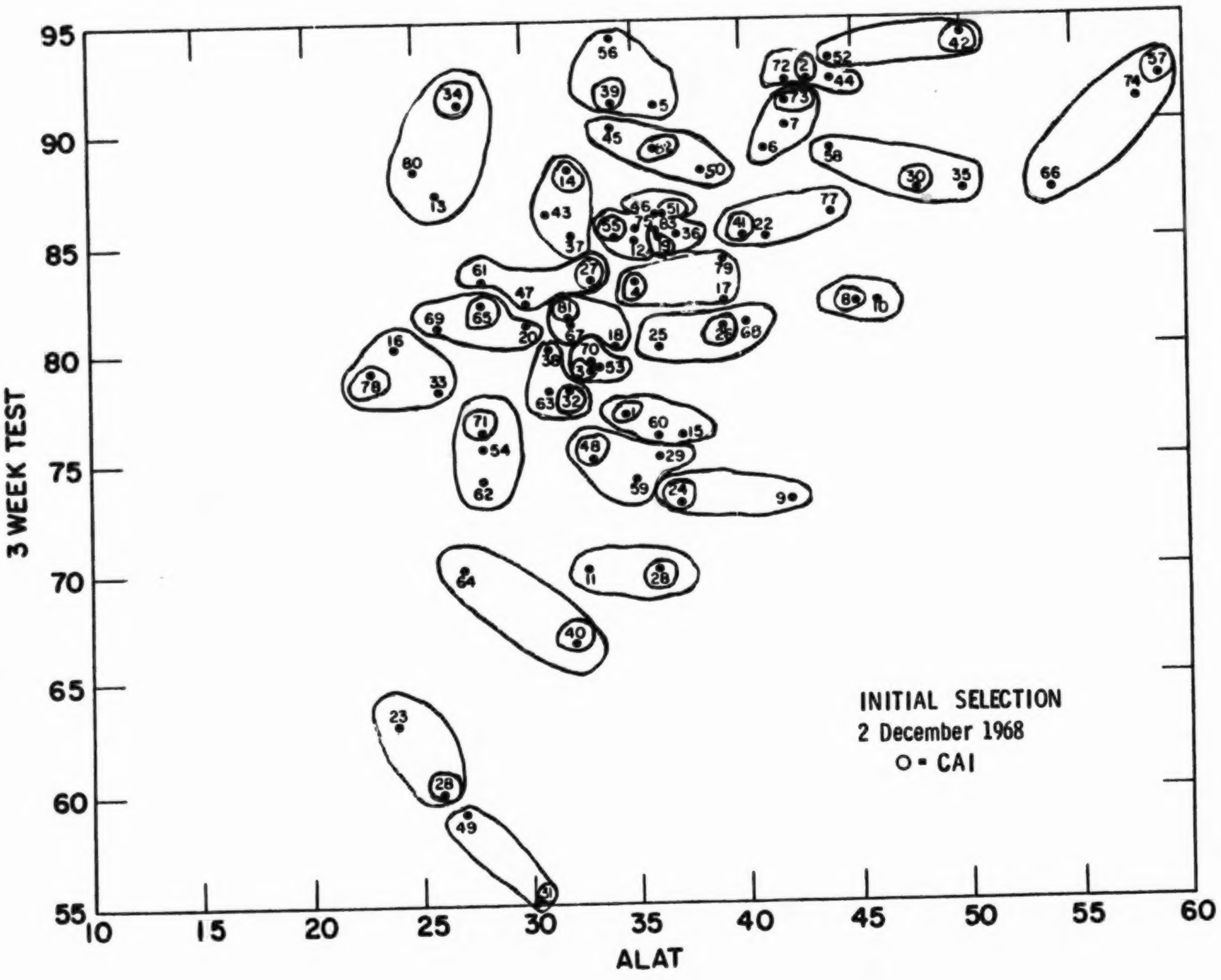
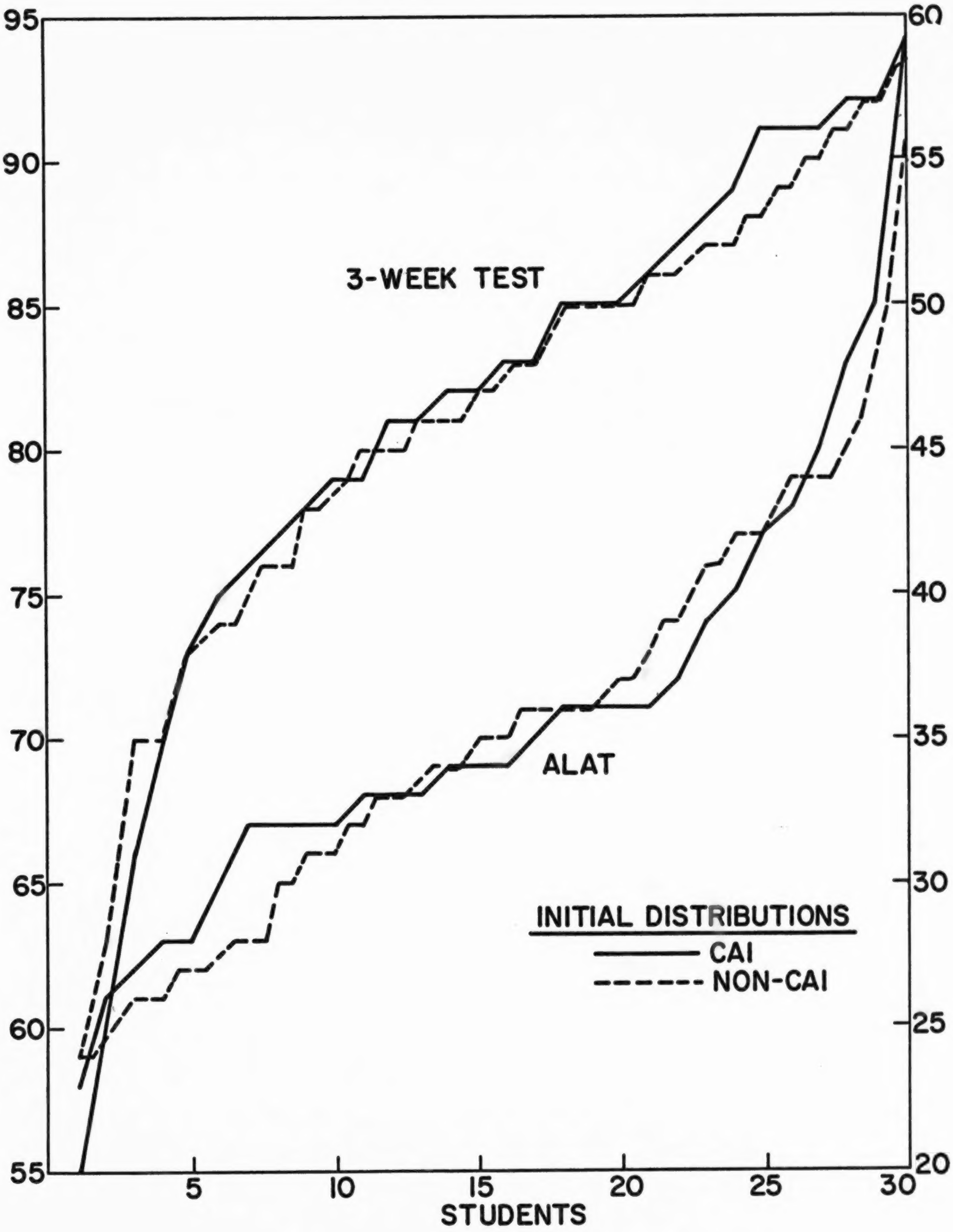

3 WEEK TEST
95
90
85
80
75
70
65
60
55
10
15
20
25
30
35
40
45
50
55
60
ALAT
INITIAL SELECTION
2 December 1968
○ = CAI

Figure II-A-3



95
90
85
80
75
70
65
60
55
60
55
50
45
40
35
30
25
20
5
10
15
20
25
30
STUDENTS
3-WEEK TEST
ALAT
INITIAL DISTRIBUTIONS
CAI
NON-CAI

Figure II-A-4

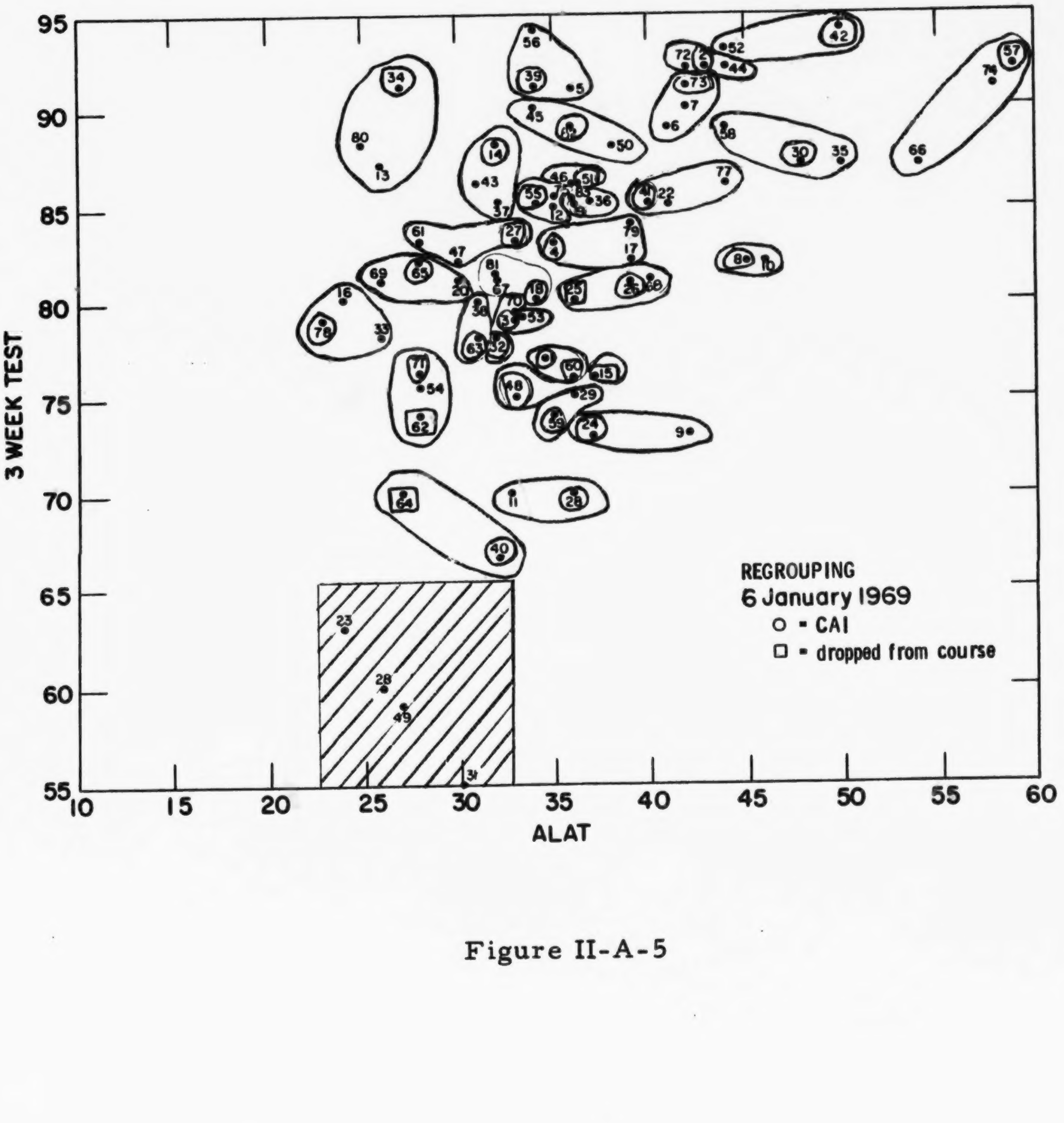

REGROUPING
6 January 1969
○ = CAI
□ = dropped from course
3 WEEK TEST
95
90
85
80
75
70
65
60
55
ALAT
10
15
20
25
30
35
40
45
50
55
60

Figure II-A-5

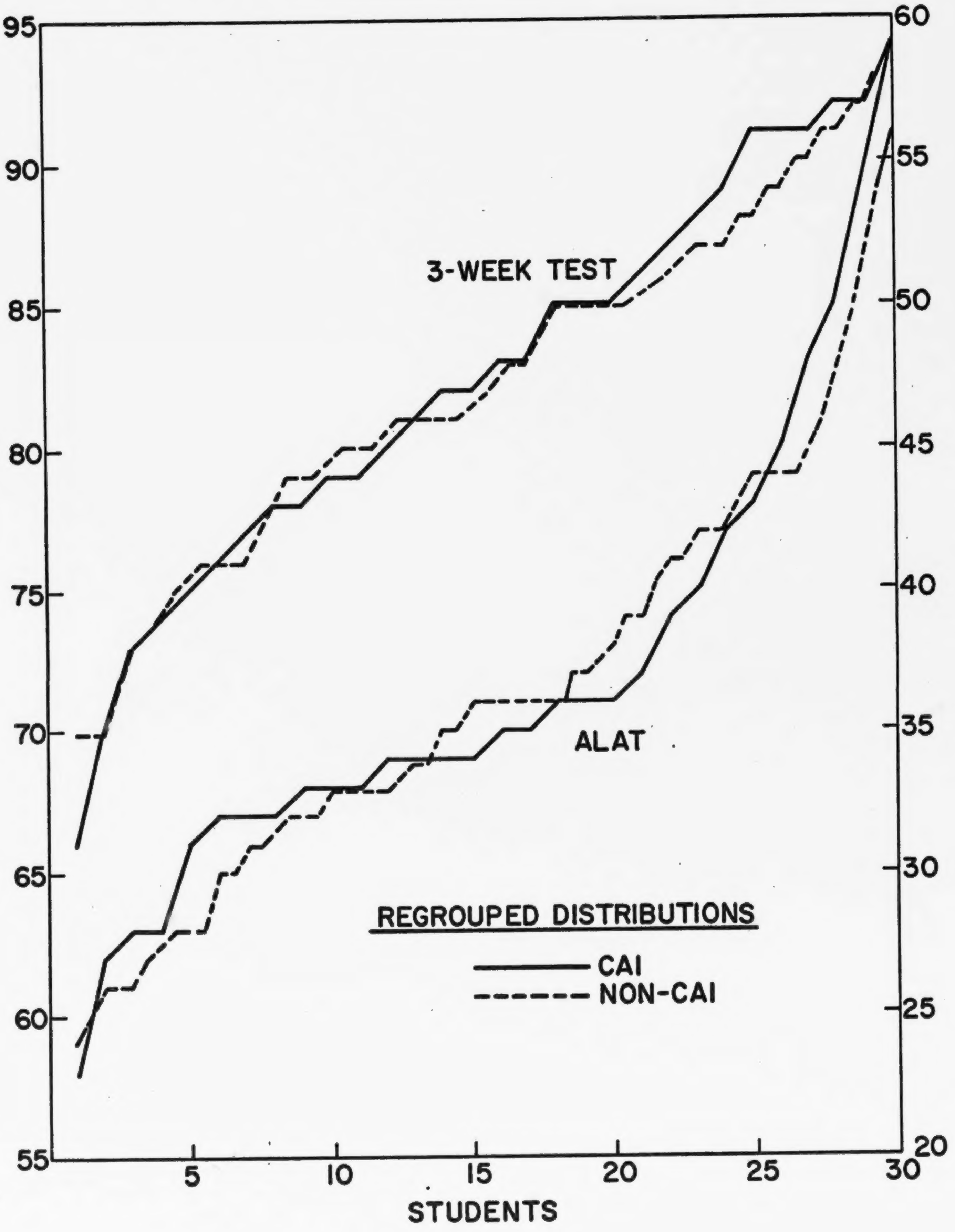

95
90
85
80
75
70
65
60
55
60
55
50
45
40
35
30
25
20
5
10
15
20
25
30
STUDENTS
3-WEEK TEST
ALAT
REGROUPED DISTRIBUTIONS
CAI
NON-CAI

Figure II-A-6

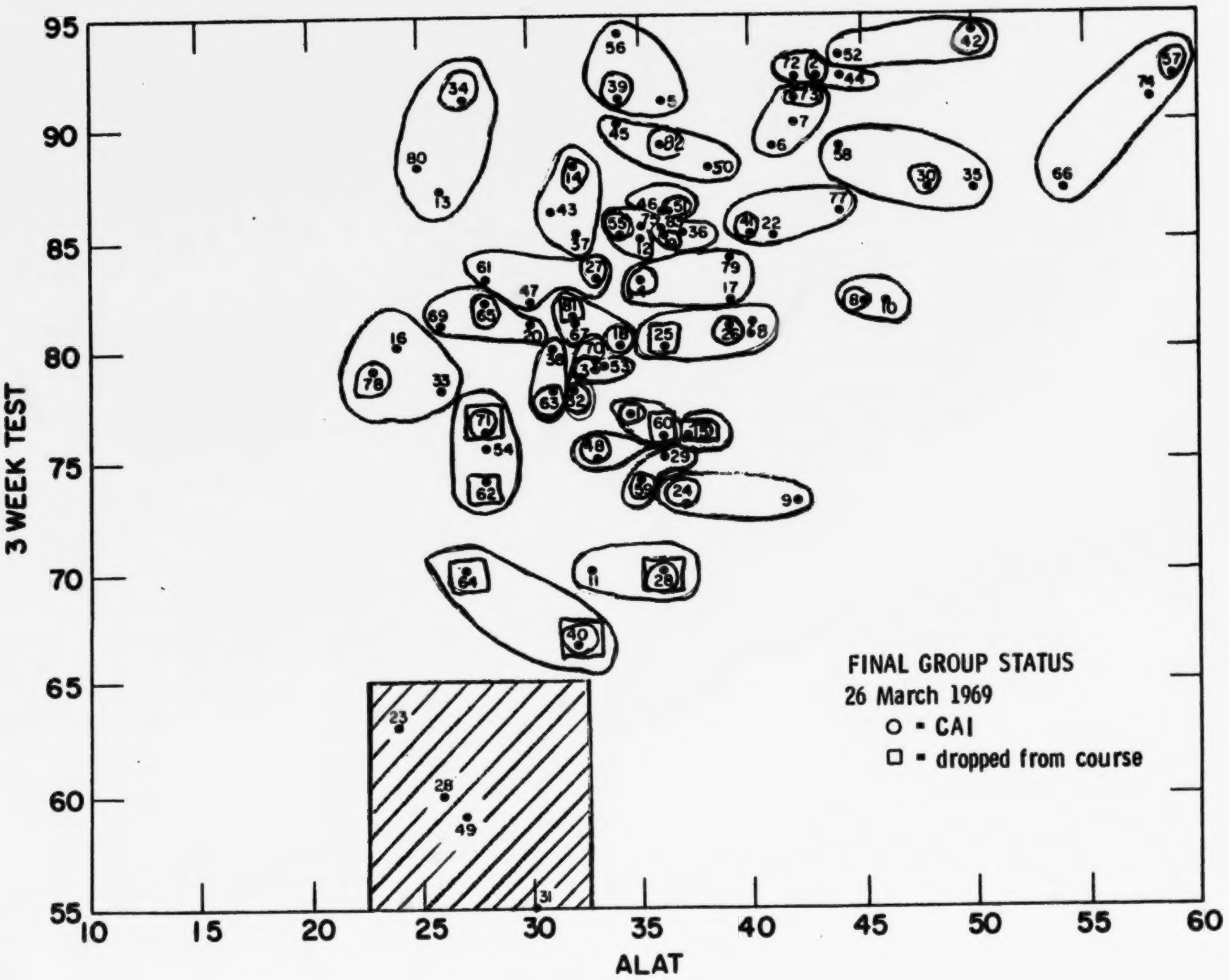

95
90
85
80
75
70
65
60
55
3 WEEK TEST
10
15
20
25
30
35
40
45
50
55
60
ALAT
FINAL GROUP STATUS
26 March 1969
○ = CAI
☐ = dropped from course
56
39
5
34
80
13
45
82
50
72
6
52
44
73
7
6
42
57
74
66
58
30
35
77
14
43
37
46
51
55
85
36
12
41
22
79
17
27
61
47
81
69
65
20
67
70
18
25
26
8
86
10
16
78
33
38
3
53
63
32
71
54
62
1
60
15
48
29
59
24
9
64
11
28
40
23
28
49
31

Figure II-A-7

Section B

This section provides an aptitude comparison, Fig. II-B-1, of CAI and non-CAI students. In constructing the aptitude comparison plot, raw ALAT scores for both sections were rank ordered, converted to raw MLAT scores, (in accordance with the conversion table II-B-2), and expressed as MLAT centiles. It is seen that the two sections are virtually identical in ALAT aptitude.

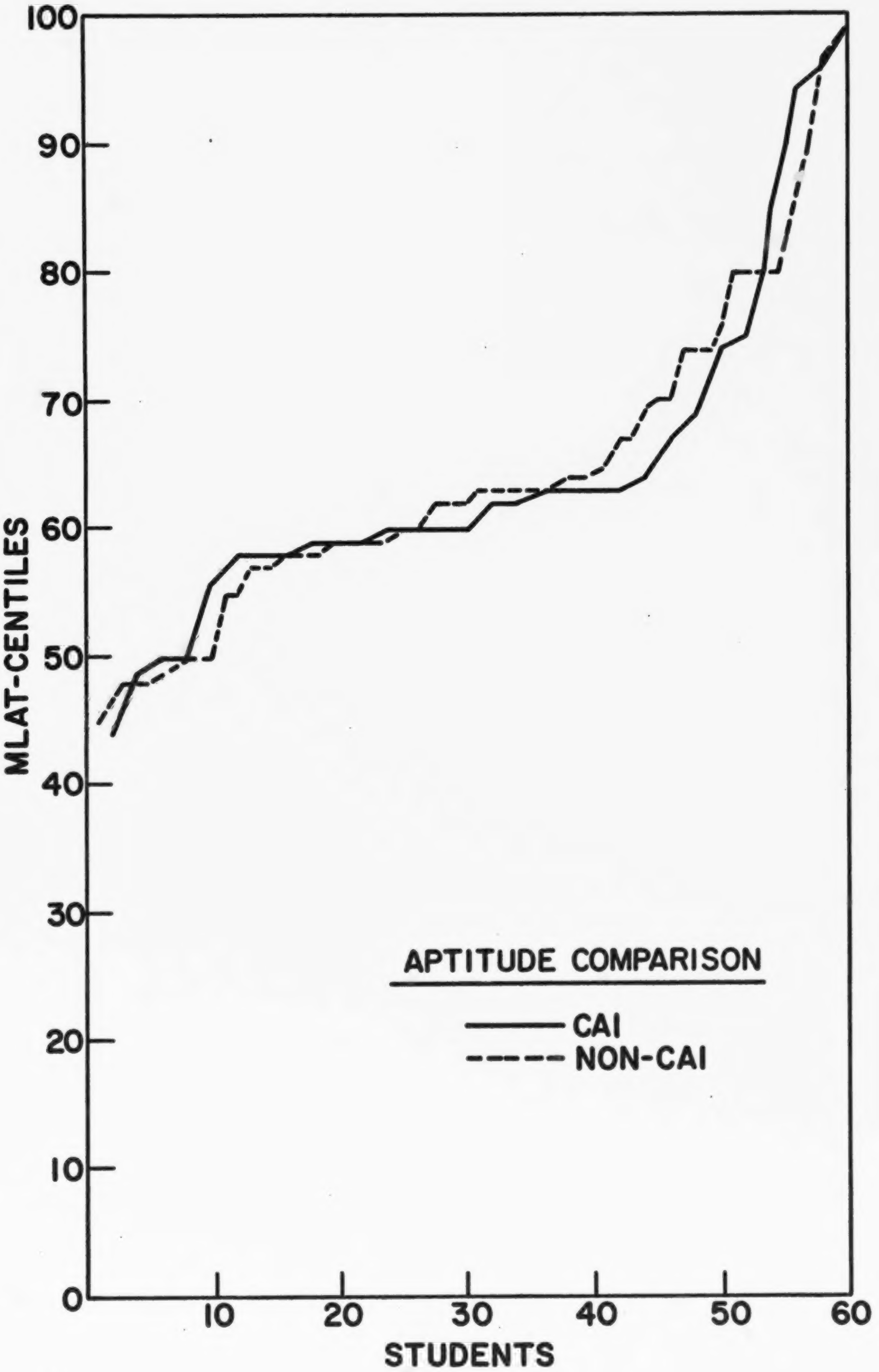

MLAT-CENTILES
0
10
20
30
40
50
60
70
80
90
100
STUDENTS
10
20
30
40
50
60
APTITUDE COMPARISON
CAI
NON-CAI

Figure II-B-1

MLAT	ALAT
162-	56-59
155-161	52-55
150-154	49-51
146-149	47-48
142-145	45-46
138-141	44
134-137	43
130-133	41-42
127-129	38-40
123-126	34-37
119-122	30-33
117-118	28-29
112-116	24-27
107-111	19-23
104-106	16-18
100-103	15
97- 99	14
91- 96	12-13
84- 90	10-11
74- 83	7- 9
57- 73	5- 6
48- 56	4
0- 47	0- 3

TABLE II-B-2 - ALAT and MLAT Raw Score Equivalence

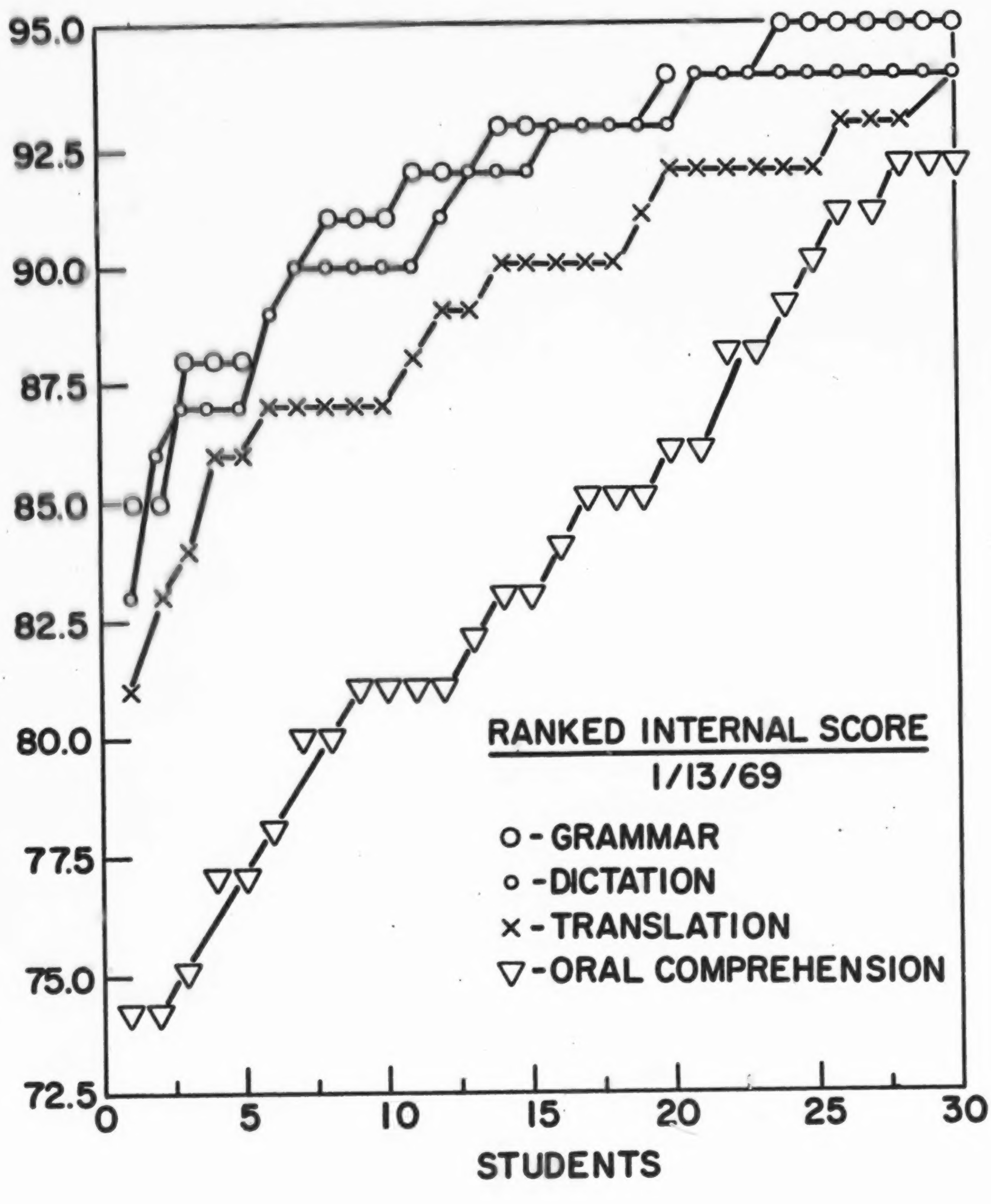

95.0
92.5
90.0
87.5
85.0
82.5
80.0
77.5
75.0
72.5
0
5
10
15
20
25
30
STUDENTS
RANKED INTERNAL SCORE
1/13/69
○ - GRAMMAR
○ - DICTATION
× - TRANSLATION
▽ - ORAL COMPREHENSION

Figure II-C-1

Appendix III

Appendix III

Evaluation

Section A

Part One

Experimental Design
for
Russian ACC Experiment in CAI
at DLI

October 9, 1968

This document states the procedural details for data collection and analysis which will be carried out in the ACC-CAI experiment at DLIWC. The five topics covered in this report are (1) the subject selection procedure, (2) the data to be collected and the data collection schedule, (3) the principal analysis, (4) supplementary analyses and their uses, and (5) suggested considerations.

The Subject Selection Procedure

The subject ACC class beginning in November will contain approximately ninety students. Of these, thirty will be selected for the experimental CAI group. A control non-CAI group will be designated from among the remaining students. Students in the CAI group will be individually matche with students in the non-CAI group on the basis of previously obtained scores on the Army Language Aptitude Test (ALAT) and the DLI 3-week test. The student selection procedure will be carried out by representatives of IBM and DLI at DLIWC on 3 December 1968. At this time, the sample population will have completed roughly three weeks of instruction and will have just taken the DLI 3-week test.

A typical DLIWC class may contain students who have a somewhat special character or status, for example, officers, WAC's, University of California students, and students who possess a native language other than English. Special attention should be paid to their distribution in the sample overall and to their distribution in the experimental and control groups.

Data to be Collected

Test Scores:

1. Scores on ALAT and 6-, 12-, and 18-week tests for at least one hundred and fifty students in ACC classes completed in the year prior to the completion of the experiment as required in the proposal (page 3-6, itema).
2. ALAT scores on all students in the November ACC class as required by the proposal (page 3-6, item b).
3. 3-week test scores for the November ACC class.
4. 6-week test scores for the November ACC class.
5. 12-week test scores for the November ACC class.
6. 18-week test scores for the November ACC class.
7. MLA Foreign Language Cooperative test scores for the November ACC class (test to be administered during the week of 24 March 1969).

CAI Performance Indices:

8. First attempt scores -- the scores achieved by students on the first attempt on an item, averaged and reported at course end.
    (a) on all items taken
    (b) on grammar drill items only (i.e., TST, ACST, ACT)
    (c) on translation items only (i.e., TRANS)
    (d) on dictation items only (i.e., DICTS, DICTNU, DICTNA)
    (e) on aural comprehension items only (i.e., ACS, ACD, IRN)
9. Module scores -- a score generated at the completion of a module based upon all attempts on all items within that module. Average module scores to be computed and reported at course end.

- (a) on all modules taken
- (b) on grammar drill modules only
- (c) on translation modules only
- (d) on dictation modules only
- (e) on aural comprehension modules only

10. Ten session recency weighted scores (taken at the time of the 6-, 12-, and 18-week tests) -- a score computed daily from the current lesson score and the current ten session recency score.
    - (a) for grammar drill only
    - (b) for translation only
    - (c) for dictation only
    - (d) for aural comprehension only

11. Items completed (daily average computed at course end)
    - (a) on all items
    - (b) on grammar drill only
    - (c) on translation only
    - (d) on dictation only
    - (e) on aural comprehension only

Data Collection Schedule:

1.	15 November 1968	(DLI to IBM)
2.	15 November 1968	(DLI to IBM)
3.	3 December 1968	(DLI to IBM)
4.	10 January 1969	(DLI to IBM)
5.	21 February 1969	(DLI to IBM)
6.	4 April 1969	(DLI to IBM)
7.	4 April 1969	(DLI to IBM)
8. - 11.	on going through operational phase	(IBM)

The Principal Analysis

A treatments-by-levels analysis of variance will be used to test for the effects of the method of instruction (CAI vs. non-CAI) and ability grouping (ALAT quartile) on each of the following criterion measures: the 18-week DLI test score, the MLA Foreign Language Cooperative sub-test scores obtained during the 18th week of instruction, and the DLI proficiency scores obtained during the 37th week of instruction. The use of this design will provide statistical analyses of (1) differences due to method of instruction, (2) the overall relationship of several measures of student performance to the ALAT, and (3) the differential effectiveness of methods of instruction at each ALAT level.

The distribution of ALAT scores will be divided into quartiles. An example of the kind of distribution that might result is illustrated in Figure 1.

		Method of Instruction		
		CAI	non-CAI	
	Q4	7	14	21
ALAT	Q3	7	14	21
LEVEL	Q2	8	16	24
	Q1	8	16	24
		30	60	

FIGURE 1.

The ALAT has been selected as the blocking variable on the basis of assurance by DLI that this measure correlates highly with important criterion measures used by DLI. The bivariate matching on the basis of ALAT and 3-week scores in the student assignment procedure should operate to maximize the similarity of distributions of the CAI and non-CAI groups on these relevant individual difference variables and thus should improve the sensitivity of the test of outcomes.

Supplementary Analyses

1. The relationships of the criterion measures to the internal indices of CAI performance (iterms 8-11 in the Data to be Collected section) will be ascertained through the use of correlational and multiple regression analyses. Interpretation and discussion of findings will be limited to instances where intercorrelations equal or exceed the .05 level of confidence (i.e., $\gamma \geq .355$). Other internal indices of performance will be monitored and, if significantly related to the criterion measures, will be reported.
2. The relations between the internal indices and the interim test scores (i.e., the 6- and 12-week tests) will be studied in order to evaluate student progress and to review the appropriateness of parameters in the CAI program.

Suggested Considerations

Should DLI consider that variability in instructor practices raises special problems, representatives of IBM and DLI should jointly review procedures for instructor selection and student testing.

Part Two

Description of Statistical Procedures and Data Processing Routines

The principal statistical procedures employed in the analysis of the data were those of correlational analysis and of analysis of variance-covariance. For the most part, these analyses were accomplished on the IBM 360/Model 91 computer by means of selected BMD Biomedial Computer Programs (University of California Press, 1968). The specific programs utilized were BMD02D (correlation with transgeneration), BMD02R (stepwise multiple regression), and BMD05V (general linear hypothesis).

The BMD02D program provided tables of intercorrelations (including means and standard deviations) among all the variables selected for investigation. BMD02R provided orthogonalization of the ALAT and three week scores. And the BMD05V program provided analysis of variance-covariance tests on each of the dependent variables for the 12 and 18 week analyses.

Since these programs are well documented and have been available through SHARE for over 7 years, a detailed description of the computational procedures will not be presented here. The interested reader should consult the BMD Program Manual. Also, since the statistical model on which the formal analyses are based-- The General Linear Hypotheses --has been well established as the appropriate procedure for designs such as that reported here, only a general description will be provided. (For example, see Scheffe, H. A., *The Analysis of Variance.* New York: Wily 1960; Winer, B. J., Statistical Principles in Experimental Design, New York: McGraw-Hill, 1962; Ostle, B., Statistics in Research, Iowa: Iowa State University Press, 1960)

For all covariance analyses, the statistical model was:

$$E(y)_{rij} = m + \alpha_i + \beta_j + \delta_{ij} + \gamma x_{ij}$$

where $i = 1,2$ (CAI vs. non)

$j = 1,2,\ldots,4$ (ALAT quartiles)

$r =$ the rth person

and

$$\sum_i \alpha_i = \sum_j \beta_j = \sum_i \delta_{ij} = \sum_j \delta_{ij} = 0$$

For each of the dependent variables, this model was evaluated for total variance explained when all terms were present as against when α_i (the treatment factor), β_j (the ALAT factor), δ_{ij} (the interaction of T x A) and γx_{ij} (the reduced 3-week covariate) were absent. The ratio: mean explained variance when all terms are present less mean explained variance when a specified term is omitted (i.e. tested) divided by mean unexplained variance when all terms are present provides an F value, which can be directly interpreted as displaying the significance (or reliablity) of the term being tested. The effects of including the covariate in the analysis on cell means is determined by evaluating the expression

$$Y_{adj} = Y_{orig} - \gamma(\overline{X}_{ij} - \overline{X}_{..})$$

$$\text{where } i = 1,2$$
$$j = 1,2,\ldots,4$$

for each variable of interest.

Section B

Tables of Grades and Scores

TABLE 1

ALAT AND 3-WEEK TEST SCORES

CAI

Student #	ALAT	3-wk. Test
1	36	89
2	43	92
3	40	85
4	48	87
5	27	91
6	34	91
7	50	94
8	36	86
9	59	92
10	42	91
11	45	82
12	33	79
13	34	77
14	36	85
15	39	81
16	33	83
17	32	88
18	34	85
19	28	82
20	35	74
21	37	73
22	35	83
23	36	70
24	31	78
25	32	78
26	32	66
27	33	75
28	34	80
29	28	76
30	23	79

non-CAI

Student #	ALAT	3-wk. Test
31	34	90
32	38	88
33	42	92
34	44	92
35	41	85
36	44	86
37	44	89
38	50	87
39	25	88
40	26	87
41	34	94
42	36	91
43	44	93
44	36	86
45	58	91
46	54	87
47	41	89
48	42	90
49	46	82
50	33	79
51	35	76
52	36	85
53	37	85
54	40	81
55	36	80
56	30	82
57	28	83
58	31	86
59	32	85
60	35	85
61	35	85
62	30	81
63	26	81
64	36	75
65	42	73
66	39	84
67	39	83
68	33	70
69	31	80
70	33	79
71	27	70
72	37	76
73	32	81
74	32	81
75	28	76
76	28	74
77	24	80
78	26	78

TABLE 2

12-WEEK GRADES

CAI

				subtests			
Student #	Overall	Class	Test	Dict./Grammar	ACS	Aural Comp.	Oral
1	83	87	78	81	80	70	80
2	88	90	85	92	80	86	83
3	90	91	89	97	91	86	83
4	85	85	85	89	86	84	82
5	88	90	85	93	89	76	82
6	94	95	93	95	99	89	88
7	93	93	92	93	88	96	89
8	88	90	85	98	85	76	81
9	96	94	97	99	99	96	92
10	91	92	90	95	97	86	83
11	77	83	71	70	65	65	82
12	80	85	74	74	71	70	82
13	80	84	76	78	69	70	86
14	84	86	81	89	75	76	82
15	78	85	70	60	76	60	82
16	81	86	76	80	67	72	83
17	82	84	79	80	78	74	84
18	86	87	84	87	86	78	84
19	81	84	78	73	77	80	82
20	83	85	81	79	76	84	85
21	72	74	69	55	76	72	72
22	73	74	72	68	69	76	75
23	65	68	61	55	53	65	70
24	80	79	81	82	82	82	76
25	72	74	70	74	69	65	70
26	62	67	57	71	50	50	65
27	72	76	65	55	56	72	74
28	75	75	74	74	71	76	75
29	61	60	62	65	61	55	68
30	72	75	68	65	70	65	70

non-CAI

				subtests			
Student #	Overall	Class	Test	Dict./Grammar	Aural Comp.	ACS	Oral
31	94	92	95	94	96	96	95
32	86	88	83	79	80	82	91
33	95	95	95	95	98	92	95
34	93	94	92	97	92	86	94
35	91	91	91	94	92	89	89
36	65	74	61	74	47	72	50
37	93	93	93	97	97	89	88
38	80	83	77	90	64	72	91
39	85	85	84	82	84	82	88
40	80	83	76	76	66	78	84
41	92	91	93	88	96	96	91
42	93	92	93	97	92	92	92
43	93	92	94	95	98	89	93
44	86	86	85	87	84	78	89
45	94	95	93	97	95	89	90
46	81	84	78	85	66	80	80
47	84	89	78	74	76	80	85
48	94	93	95	92	96	96	96
49	87	87	87	83	83	89	90
50	82	83	80	88	72	78	80
51	-	-	-	-	-	-	-
52	86	86	85	83	91	82	82
53	84	83	85	88	82	70	90
54	75	76	73	90	59	70	74
55	-	-	-	-	-	-	-
56	80	80	80	70	78	80	91
57	86	87	84	87	86	78	86
58	82	83	80	77	70	84	90
59	88	87	88	81	84	92	95
60	89	89	88	82	88	89	92
61	75	76	74	65	74	78	80
62	83	81	84	90	88	82	77
63	71	70	71	74	57	72	79
64	68	68	68	65	78	70	60
65	70	71	67	75	55	78	60
66	81	80	82	78	89	80	80
67	75	76	74	80	82	65	70
68	72	72	71	65	81	74	65
69	-	-	-	-	-	-	-
70	76	75	77	75	86	70	78
71	-	-	-	-	-	-	-
72	-	-	-	-	-	-	-
73	81	80	82	77	78	84	90
74	78	75	81	74	79	84	85
75	80	76	84	86	81	84	85
76	-	-	-	-	-	-	-
77	79	79	79	82	81	78	76
78	73	74	72	74	69	72	72

TABLE 3

12-WEEK MLA (Form MA) SUBTESTS

CAI

Student #	Reading	Writing	Speaking	Listening
1	23	38	20	21
2	21	46	35	18
3	25	56	40	22
4	21	54	32	16
5	21	38	37	21
6	31	59	52	30
7	35	63	45	30
8	15	50	33	24
9	25	69	50	29
10	19	53	42	32
11	17	24	25	18
12	17	37	29	18
13	9	40	29	16
14	11	38	30	14
15	14	24	20	15
16	21	35	26	20
17	18	35	24	19
18	16	31	36	20
19	17	27	25	23
20	21	34	35	20
21	19	19	21	10
22	20	26	25	18
23	22	16	23	17
24	11	39	26	22
25	17	35	19	15
26	15	21	-	14
27	16	30	27	13
28	19	28	26	15
29	11	19	-	11
30	19	21	30	21

non-CAI

Student #	Reading	Writing	Speaking	Listening
31	18	50	50	29
32	25	37	22	22
33	22	56	42	29
34	29	48	33	30
35	27	54	44	22
36	16	41	21	16
37	22	52	34	25
38	19	42	34	17
39	6	40	24	14
40	19	33	31	21
41	16	45	40	27
42	18	52	26	19
43	18	57	47	26
44	15	46	28	17
45	33	62	47	27
46	20	46	23	13
47	19	40	33	23
48	28	59	50	27
49	26	45	30	18
50	22	52	27	19
51	-	-	-	-
52	18	36	30	21
53	18	36	27	22
54	18	47	24	17
55	-	-	-	-
56	16	33	27	22
57	18	39	37	21
58	17	37	31	13
59	29	41	31	31
60	14	42	34	23
61	20	21	22	21
62	10	38	30	23
63	10	22	[illegible]	16
64	23	17	29	19
65	17	29	21	21
66	25	26	30	23
67	20	34	28	19
68	21	30	28	15
69	-	-	-	-
70	19	50	20	25
71	-	-	-	-
72	-	-	-	-
73	18	35	28	27
74	18	33	42	18
75	23	36	20	21
76	-	-	-	-
77	16	36	29	19
78	21	35	20	23

TABLE 4

18-WEEK GRADES

CAI

				subtests				
Student #	Overall	Class	Test	Aural Comp. Sent.	Aural Comp.	Dict.	Grammar	Oral
1	80	82	78	88	78	77	72	75
2	86	86	86	85	80	82	92	88
3	92	92	92	86	88	98	97	83
4	85	85	85	87	84	75	95	84
5	91	91	91	89	92	88	95	87
6	96	96	95	90	100	92	95	96
7	95	94	95	100	88	98	95	91
8	89	89	89	81	100	83	96	79
9	95	96	93	100	92	84	96	93
10	90	92	87	87	96	81	86	87
11	75	73	76	81	70	74	75	80
12	79	85	73	75	61	61	88	76
13	82	84	80	75	78	72	92	80
14	82	85	78	70	78	80	82	76
15	76	75	76	80	74	60	84	82
16	81	83	78	73	74	73	82	85
17	81	83	78	73	72	75	86	76
18	89	89	88	84	92	85	92	81
19	81	82	79	71	96	80	74	77
20	83	86	80	69	74	82	76	79
21	72	73	70	77	78	60	62	76
22	77	75	78	74	80	76	79	79
23	-	-	-	-	-	-	-	-
24	84	85	82	86	84	88	75	77
25	71	75	66	68	72	50	71	73
26	-	-	-	-	-	-	-	-
27	73	76	70	69	66	62	76	79
28	75	76	73	64	72	64	79	85
29	-	-	-	-	-	-	-	-
30	70	72	68	73	74	58	64	80

non-CAI

				subtests				
Student #	Overall	Class	Test	Aural Comp. Sent.	Aural Comp.	Dict.	Grammar	Oral
31	92	92	92	87	98	92	94	89
32	87	87	86	83	74	90	90	87
33	95	95	95	97	88	96	98	91
34	92	92	92	93	96	84	98	91
35	92	92	92	83	98	94	92	89
36	73	74	72	73	70	71	72	72
37	93	94	92	83	98	94	94	90
38	83	85	80	75	68	85	86	79
39	86	87	85	90	88	86	80	84
40	83	86	80	75	80	80	86	76
41	92	90	94	91	96	94	96	88
42	90	91	89	89	84	82	98	92
43	90	92	88	83	88	84	94	86
44	87	84	89	95	80	90	98	78
45	96	95	96	94	96	98	96	93
46	84	85	83	77	78	83	88	82
47	87	86	88	88	92	84	94	80
48	95	93	96	95	96	96	96	94
49	84	85	83	90	92	79	77	81
50	84	85	82	78	74	77	96	78
51	-	-	-	-	-	-	-	-
52	88	87	89	92	92	94	82	86
53	84	85	82	74	78	80	97	74
54	81	80	82	66	84	78	98	75
55	-	-	-	-	-	-	-	-
56	83	82	84	81	78	84	93	78
57	87	85	88	84	80	90	90	90
58	83	83	83	83	74	79	93	84
59	81	79	83	83	84	87	77	85
60	88	84	92	88	92	98	90	86
61	72	66	77	78	76	73	84	71
62	87	86	87	81	88	87	90	87
63	72	71	72	64	70	74	75	74
64	71	68	73	68	84	76	72	65
65	78	75	80	69	78	76	94	73
66	80	80	80	83	76	75	82	80
67	79	77	80	89	76	82	80	72
68	72	69	74	76	74	71	79	65
69	-	-	-	-	-	-	-	-
70	78	73	82	83	78	85	82	77
71	-	-	-	-	-	-	-	-
72	-	-	-	-	-	-	-	-
73	83	81	85	83	84	82	92	81
74	60	60	60	60	60	60	60	60
75	81	77	85	82	88	84	88	78
76	-	-	-	-	-	-	-	-
77	79	78	79	87	80	74	80	76
78	79	76	81	72	78	83	94	73

TABLE 5

18-Week MLA (Form MB) SCORES

CAI

Student #	Reading	Writing	Speaking	Listening
1	16	61	34	22
2	25	55	47	24
3	24	78	53	30
4	21	70	38	25
5	22	61	51	27
6	32	69	52	33
7	27	67	60	35
8	20	68	40	25
9	36	76	60	37
10	18	65	54	31
11	13	36	42	20
12	19	57	35	20
13	18	48	27	18
14	21	60	46	20
15	21	43	26	16
16	17	51	28	23
17	16	63	34	26
18	21	61	46	28
19	18	52	30	19
20	19	68	50	29
21	16	46	30	22
22	14	44	30	15
23	-	-	-	-
24	14	61	36	27
25	19	40	28	18
26	-	-	-	-
27	15	45	26	17
28	18	39	31	23
29	-	-	-	-
30	18	34	24	22

non-CAI

Student #	Listening	Speaking	Writing	Reading
31	31	57	50	34
32	27	71	51	26
33	29	78	56	35
34	33	78	56	35
35	34	52	48	32
36	24	57	36	15
37	20	75	40	33
38	21	58	37	23
39	7	61	43	24
40	21	52	42	25
41	19	69	46	33
42	19	60	44	37
43	20	75	36	30
44	23	71	36	28
45	35	83	51	35
46	19	72	33	26
47	24	52	43	30
48	24	78	57	35
49	24	53	45	32
50	18	55	35	25
51	-	-	-	-
52	17	57	43	29
53	25	47	37	19
54	23	64	32	21
55	-	-	-	-
56	16	41	36	21
57	20	64	49	30
58	10	57	40	27
59	24	52	52	32
60	19	69	42	31
61	18	35	27	22
62	24	69	50	33
63	18	36	29	21
64	16	40	42	24
65	20	51	44	23
66	26	43	37	30
67	21	57	39	22
68	19	45	29	27
69	-	-	-	-
70	14	50	32	15
71	-	-	-	=
72	-	-	-	-
73	27	46	38	28
74	-	-	-	-
75	20	68	35	29
76	-	-	-	-
77	22	38	39	24
78	13	54	39	26

TABLE 6

12-WEEK COMPUTER GENERATED SCORES

Student #	Grammar	Translation	Dictation	Aural Comprehension	Average Time on System
1	91.5	86.5	90	88	79.8
2	91	86	91.5	87.5	78.8
3	93.5	90.5	93.5	90	63
4	92	92	92	94.5	70
5	91.5	86	93	91	79.6
6	93	90.5	93.5	90.5	52.4
7	95	94	93.5	93	53.3
8	92.5	89	92.5	87.5	66.4
9	94.5	91.5	93.5	94.5	51.8
10	93	90.5	93	92	70.3
11	83	78.5	86	82	90
12	90.5	88.5	90	90.5	80.7
13	88.5	87	87	88	83
14	90	85	89.5	86.5	74.9
15	88.5	84.5	85.5	85.5	89.6
16	87.5	86.5	88	87	80.2
17	87.5	85.5	89	87.5	83.2
18	91	87	91.5	90.5	78.5
19	87.5	83.5	85	86	80.1
20	89.5	86	90.5	88	86
21	83	83.5	83.5	84.5	86.8
22	89	81	88.5	82.5	92.9
23	86	80	84.5	82	96.1
24	93	90	92	88	89.2
25	91	86.5	83	85	92.9
26	87.5	83.5	85	80	96.3
27	87.5	84.5	81	81	96.8
28	83.5	82.5	80	78.5	92.9
29	83	78	84.5	82.5	98
30	88.5	84	85.5	87.5	91.8

4

TABLE 7

18-WEEK COMPUTER GENERATED SCORES

Student #	Grammar	Translation	Dictation	Aural Comprehension
1	91	86	90	90
2	90	85	98	96
3	96	95	96	96
4	98	96	97	98
5	95	92	97	96
6	98	94	98	94
7	96	94	96	97
8	91	89	96	91
9	97	95	98	96
10	94	93	95	94
11	80	79	90	86
12	95	92	95	93
13	91	87	92	90
14	93	88	93	88
15	91	87	92	91
16	90	88	89	92
17	95	95	96	98
18	93	90	95	95
19	87	85	87	93
20	95	89	96	90
21	88	82	86	88
22	89	84	91	89
23	90	82	93	82
24	91	91	95	91
25	97	77	94	86
26	85	82	85	78
27	86	84	89	88
28	85	78	89	88
29	83	78	83	82
30	88	81	89	90

TABLE 8

CAI--Average Time on System at 18 Weeks

(survivors only)

1	0.79
2	0.74
3	0.62
4	0.69
5	0.74
6	0.56
7	0.54
8	0.69
9	0.52
10	0.7
11	0.86
12	0.77
13	0.82
14	0.75
15	0.89
16	0.82
17	0.83
18	0.76
19	0.82
20	0.81
21	0.82
22	0.88
23	0.84
24	0.92
25	0.97
26	0.95
27	0.88

TABLE 10

24-WEEK GRADES

CAI

Student #	Overall	Class	Test	subtests Oral	Aural Comp.	Grammar	Dict.
1	-	-	-	-	-	-	-
2	88	88	88	85	93	90	82
3	94	93	95	97	97	90	95
4	78	76	79	75	81	84	76
5	91	89	92	92	94	95	88
6	95	95	94	95	91	95	94
7	96	95	97	99	98	95	95
8	85	84	85	78	93	88	81
9	98	97	99	100	100	98	96
10	91	87	94	92	96	95	92
11	78	77	78	79	75	76	80
12	83	78	88	88	87	93	82
13	78	77	78	75	78	81	78
14	85	85	84	88	85	81	81
15	76	75	77	71	80	81	76
16	77	78	76	83	74	69	76
17	77	77	77	78	77	79	75
18	88	87	89	90	91	84	89
19	75	76	74	74	69	81	70
20	88	85	90	91	81	93	84
21	68	66	69	55	72	81	77
22	76	74	78	89	67	81	76
23	-	-	-	-	-	-	-
24	87	87	86	90	92	86	77
25	75	76	74	75	79	63	79
26	-	-	-	-	-	-	-
27	-	-	-	-	-	-	-
28	78	77	79	81	82	69	85
29	-	-	-	-	-	-	-
30	69	69	68	74	62	65	70

non-CAI

Student #	Overall	Class	Test	subtests Oral	Aural Comp.	Grammar	Dict.
31	94	94	93	94	95	90	93
32	91	88	93	95	96	90	90
33	97	96	97	99	98	91	98
34	96	95	96	96	96	95	97
35	91	91	91	97	87	91	87
36	79	77	80	90	76	79	73
37	95	94	96	99	95	93	95
38	84	86	81	76	81	83	82
39	89	88	90	96	91	93	80
40	84	84	83	85	84	81	80
41	87	86	87	91	83	91	81
42	92	90	93	91	96	96	90
43	91	89	92	94	93	95	84
44	87	87	87	88	90	79	89
45	97	97	96	95	99	93	93
46	84	84	84	88	79	83	85
47	83	84	82	84	75	95	84
48	96	95	96	94	95	96	97
49	87	85	89	89	86	91	88
50	85	86	83	89	81	88	75
51	-	-	-	-	-	-	-
52	90	88	92	95	95	90	88
53	85	84	85	80	88	91	79
54	86	85	86	88	99	79	78
55	-	-	-	-	-	-	-
56	76	77	74	69	67	83	74
57	87	86	87	88	89	84	86
58	81	83	79	81	77	84	74
59	78	78	78	73	72	93	72
60	88	89	87	94	89	83	82
61	73	69	76	75	72	83	74
62	89	88	89	88	96	90	81
63	74	75	72	76	71	69	73
64	65	66	63	57	51	74	70
65	85	84	86	91	85	88	80
66	79	80	78	86	70	79	75
67	82	83	81	91	79	79	73
68	76	74	77	83	75	81	70
69	-	-	-	-	-	-	-
70	80	79	80	81	78	86	75
71	-	-	-	-	-	-	-
72	-	-	-	-	-	-	-
73	84	82	86	92	81	91	80
74	-	-	-	-	-	-	-
75	85	85	84	89	81	83	80
76	-	-	-	-	-	-	-
77	86	84	87	91	87	86	85
78	81	83	78	78	73	81	78

Section C

Section C contains performance comparisons for CAI and non-CAI sections on the tests and sub-tests administered after 12 weeks of instruction. As in the predictor variable comparisons illustrated in Appendix II, Section A, the points plotted are rank order raw scores with non-CAI students weighted either one or one-half depending on whether they originate from groups containing one or two non-CAI students respectively.

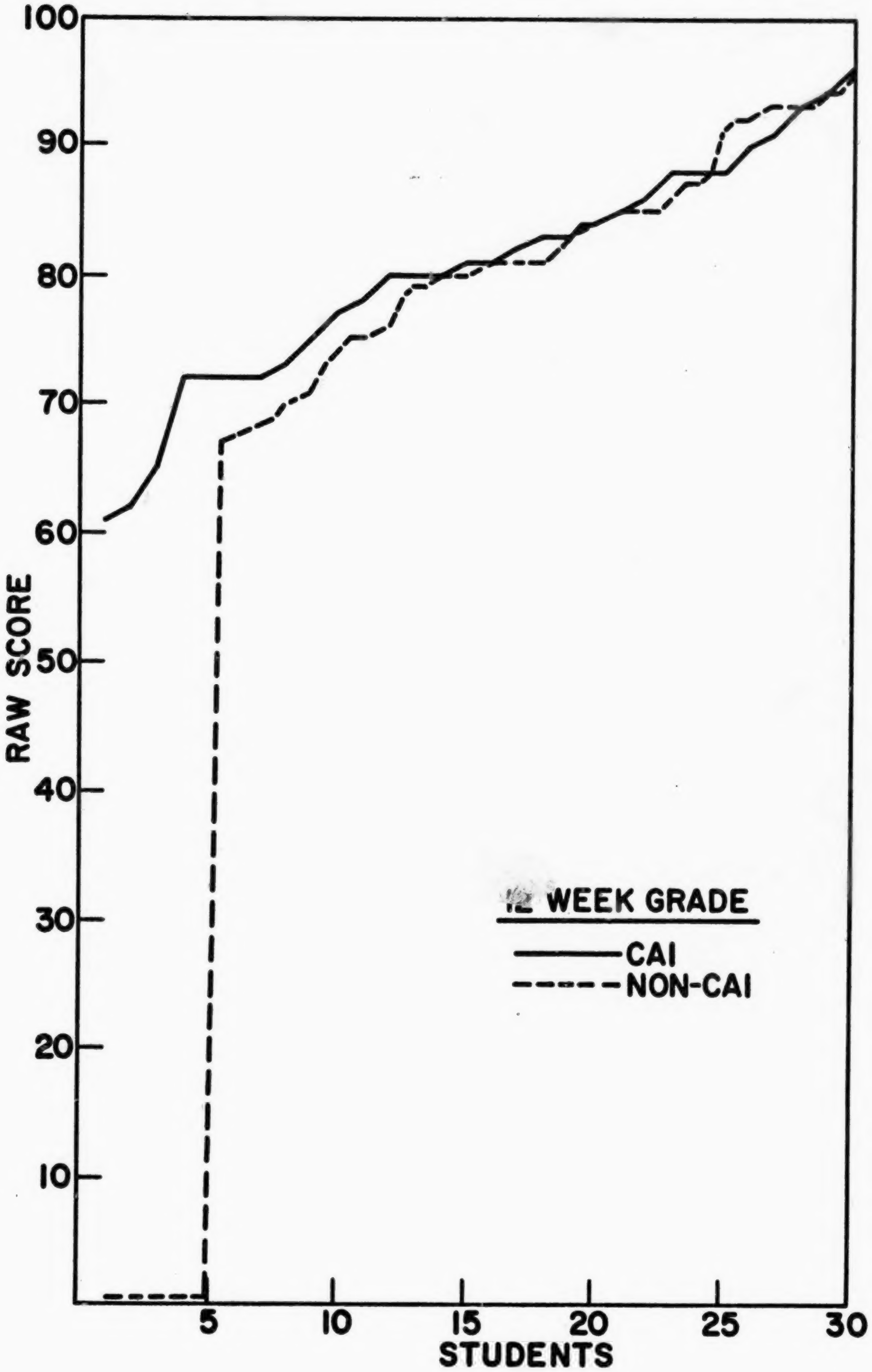


Figure III-C-1

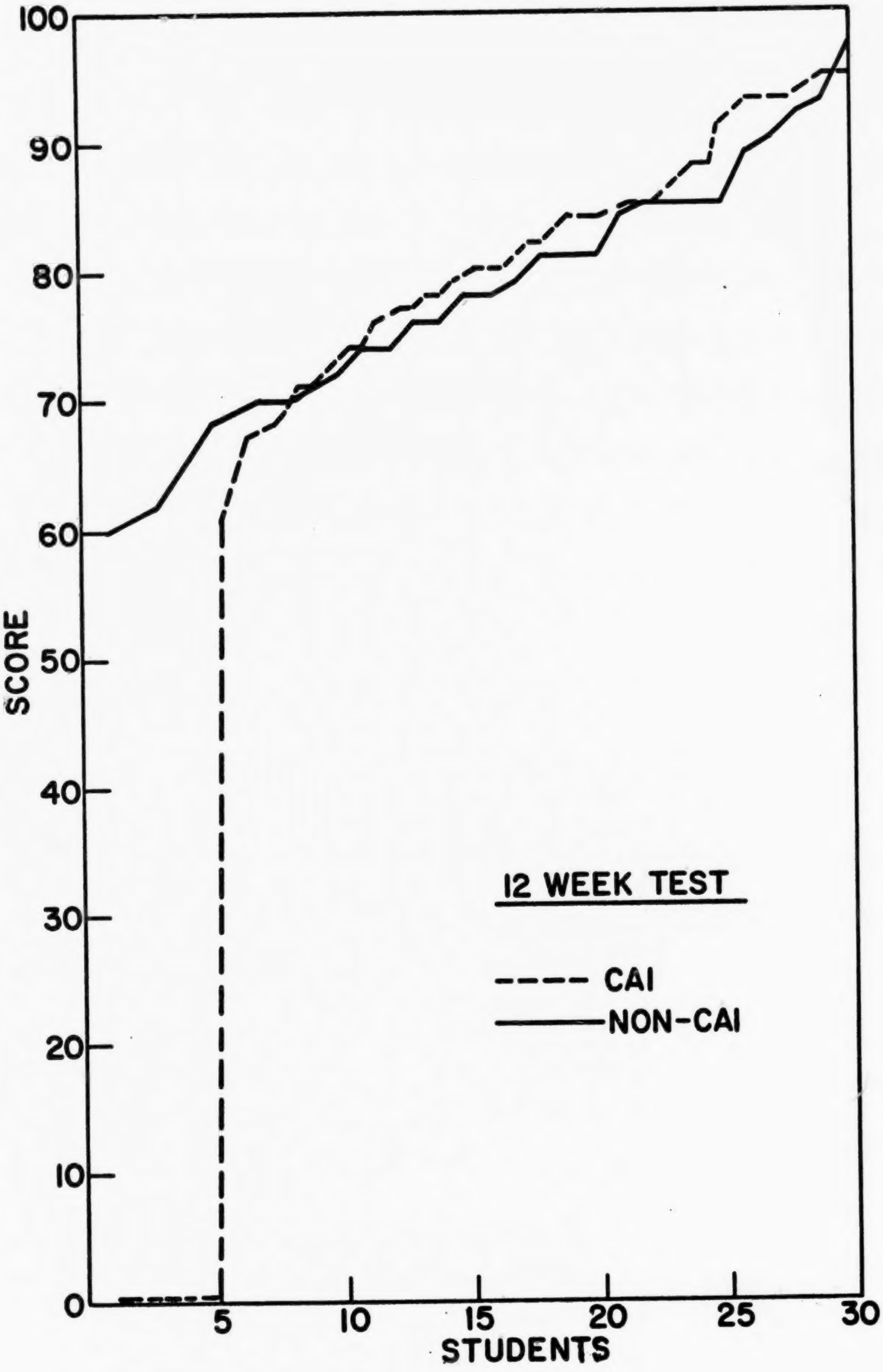


Figure III-C-2

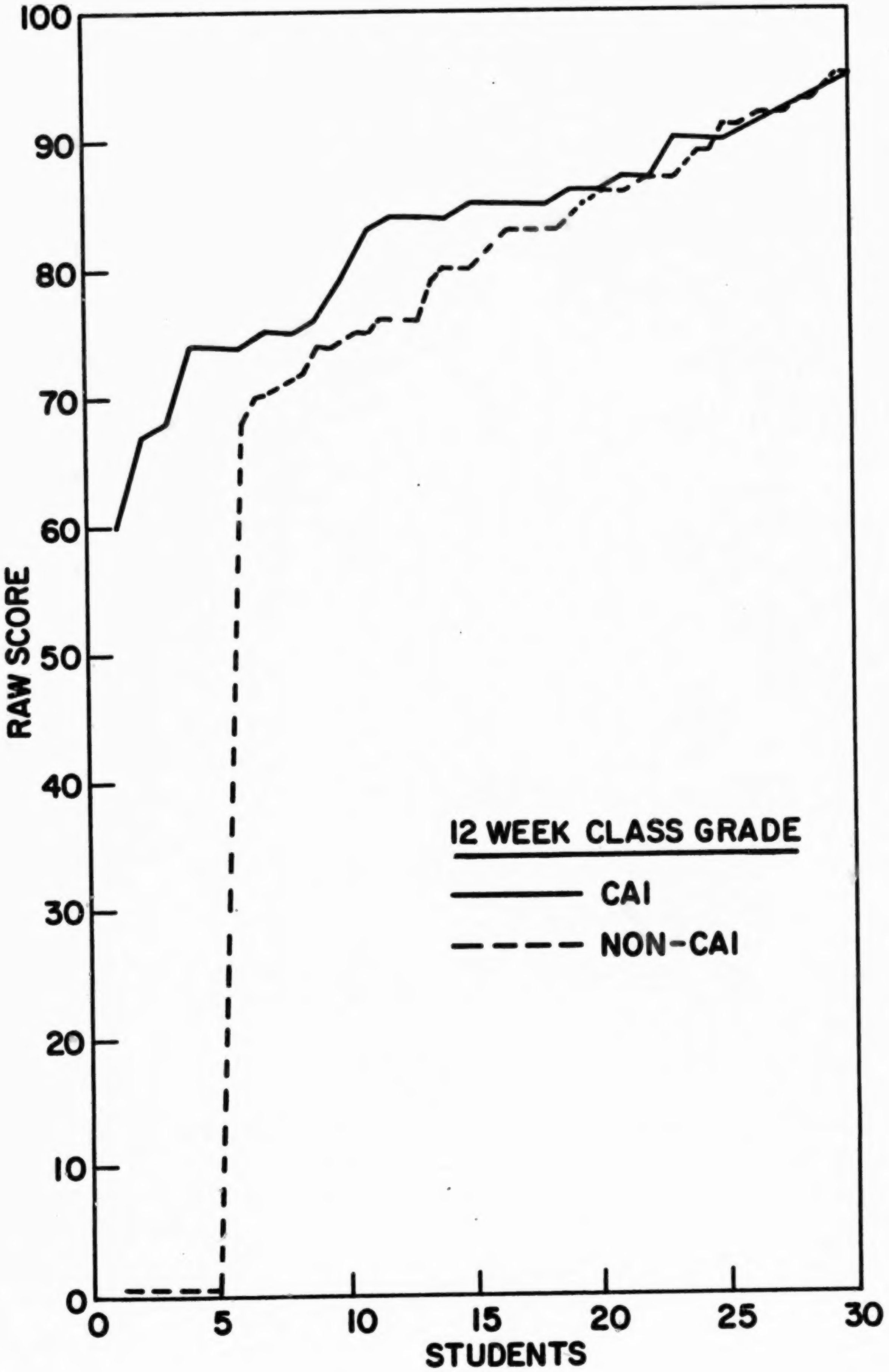


Figure III-C-3

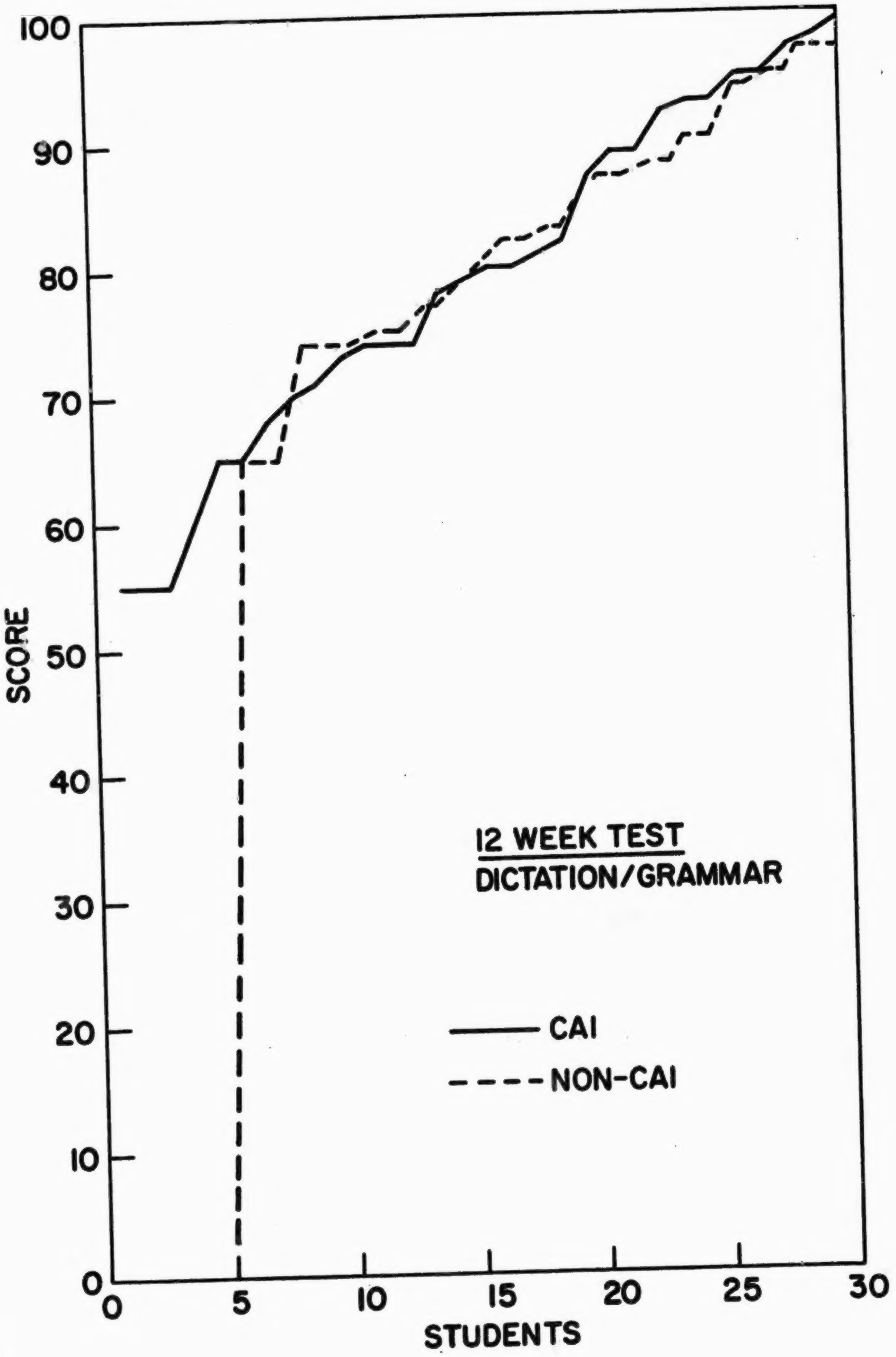


Figure III-C-4

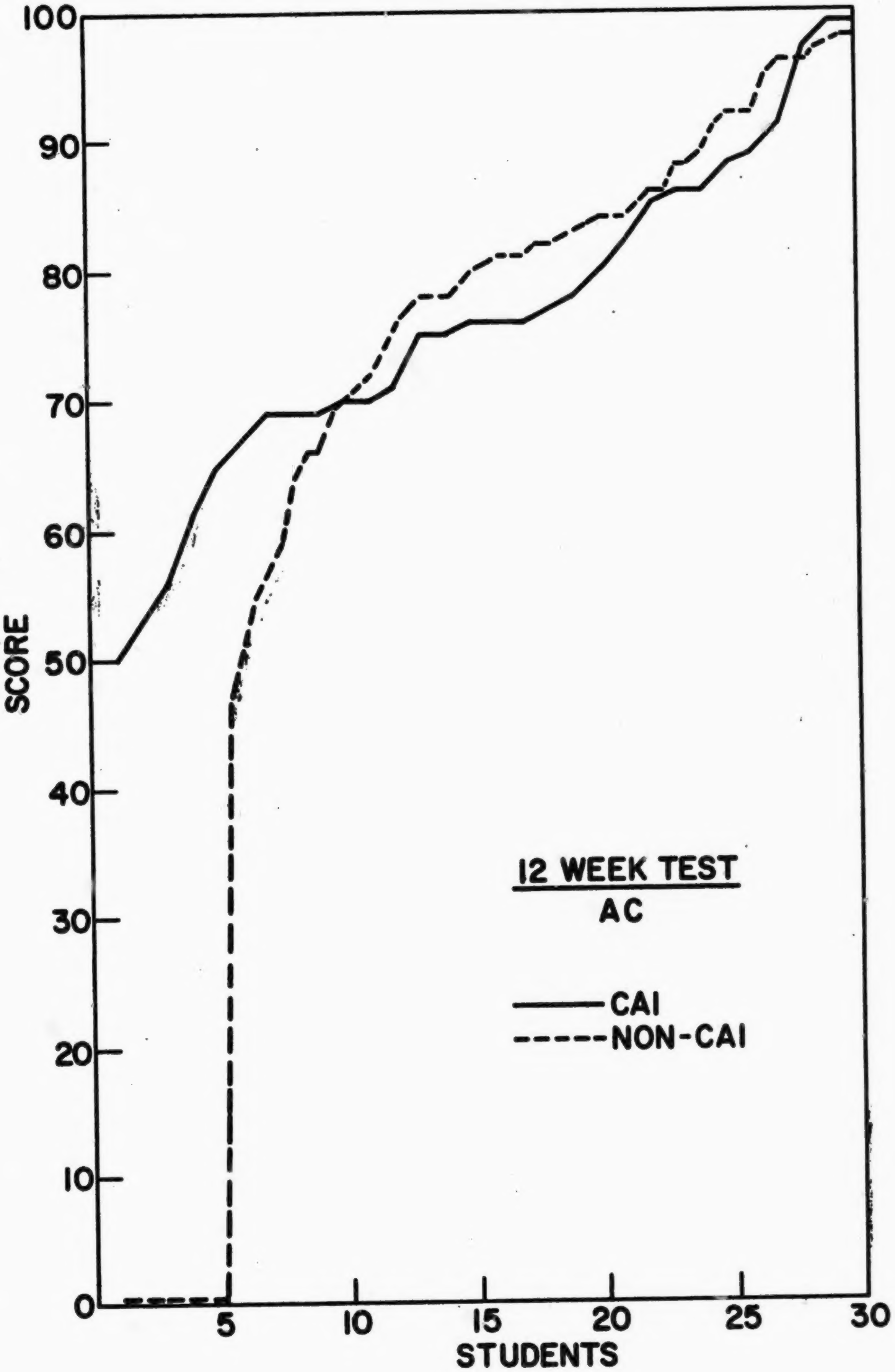


Figure III-C-5

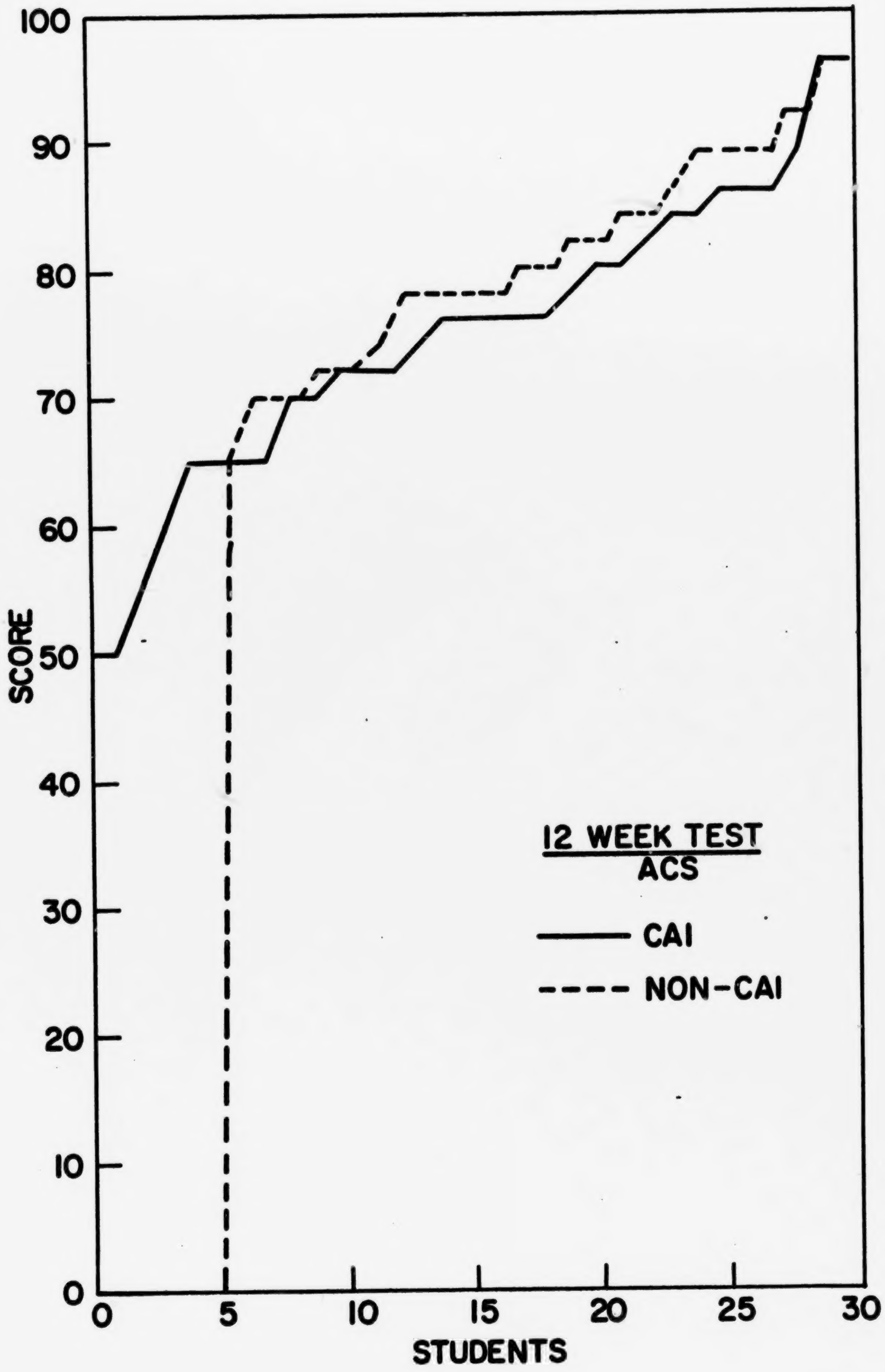


Figure III-C-6

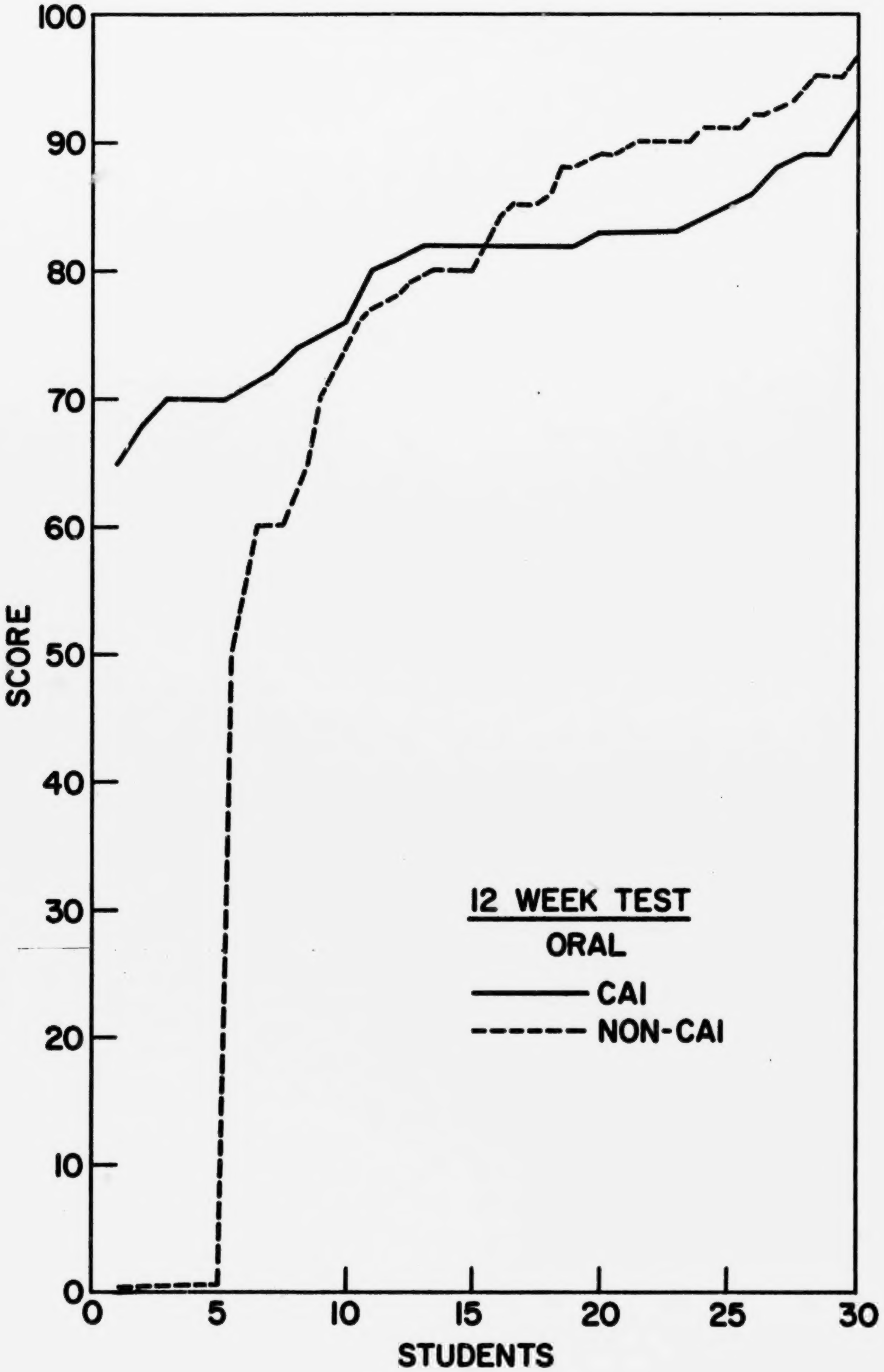


Figure III-C-7

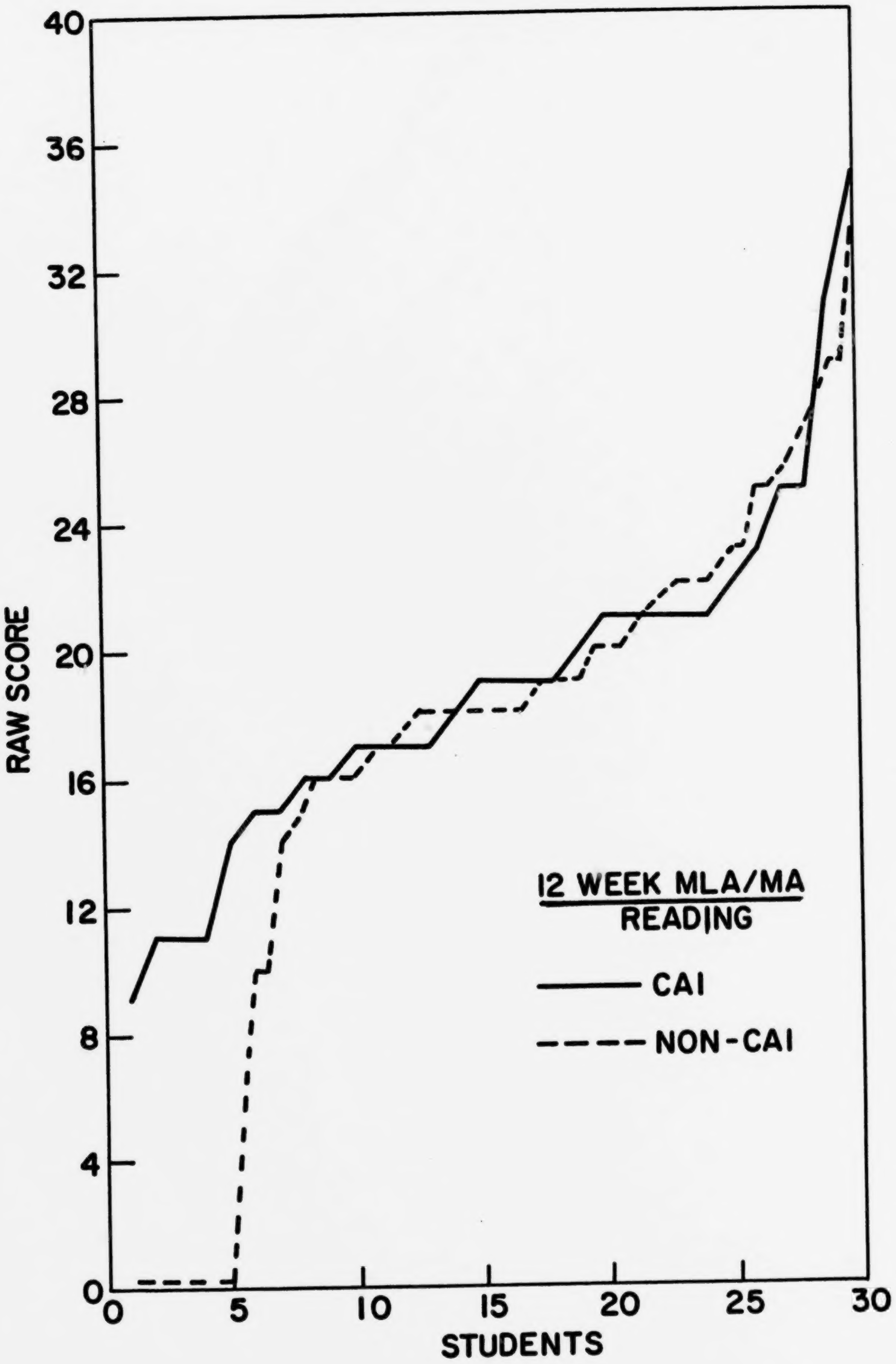


Figure III-C-8

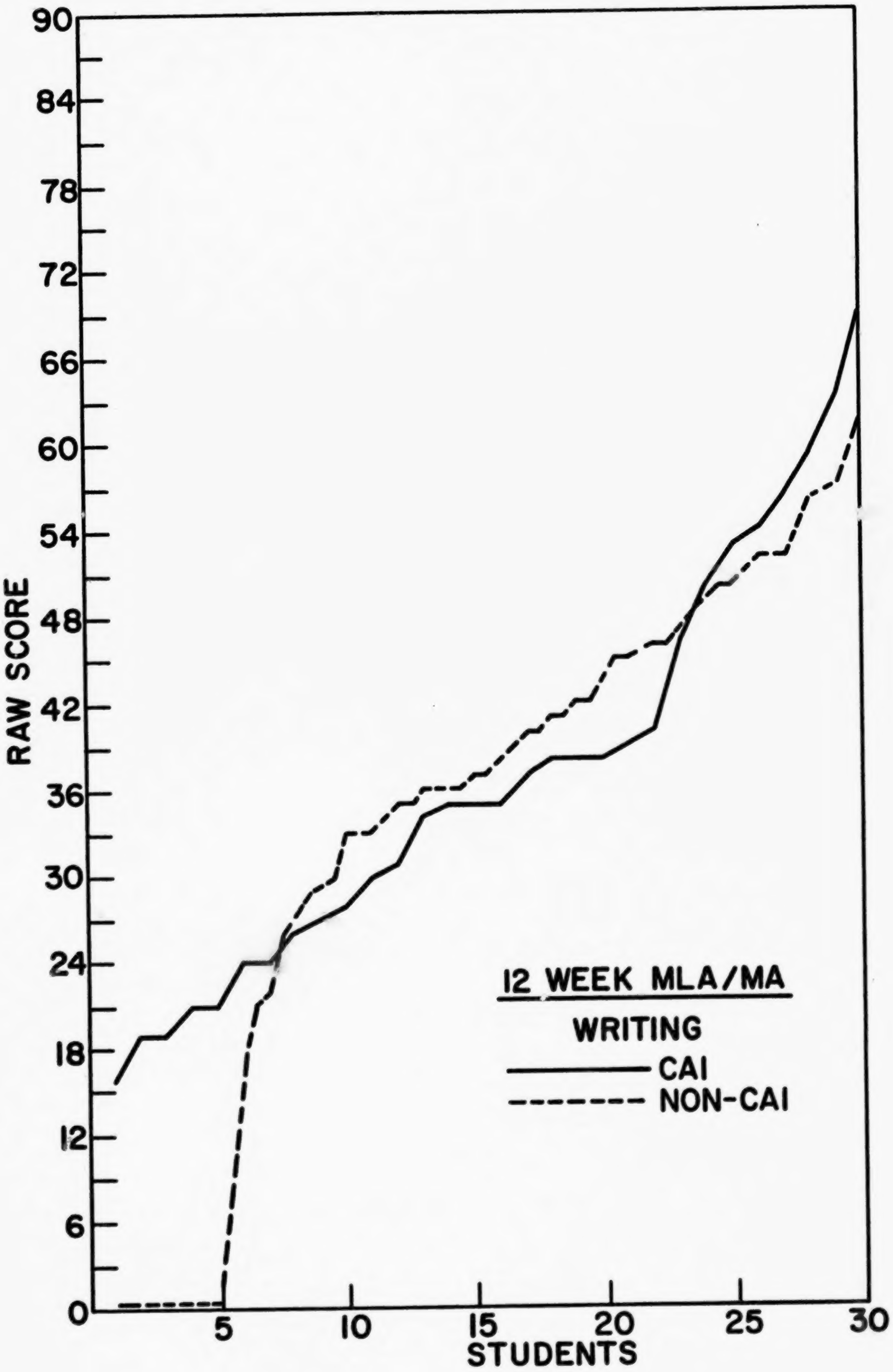


Figure III-C-9

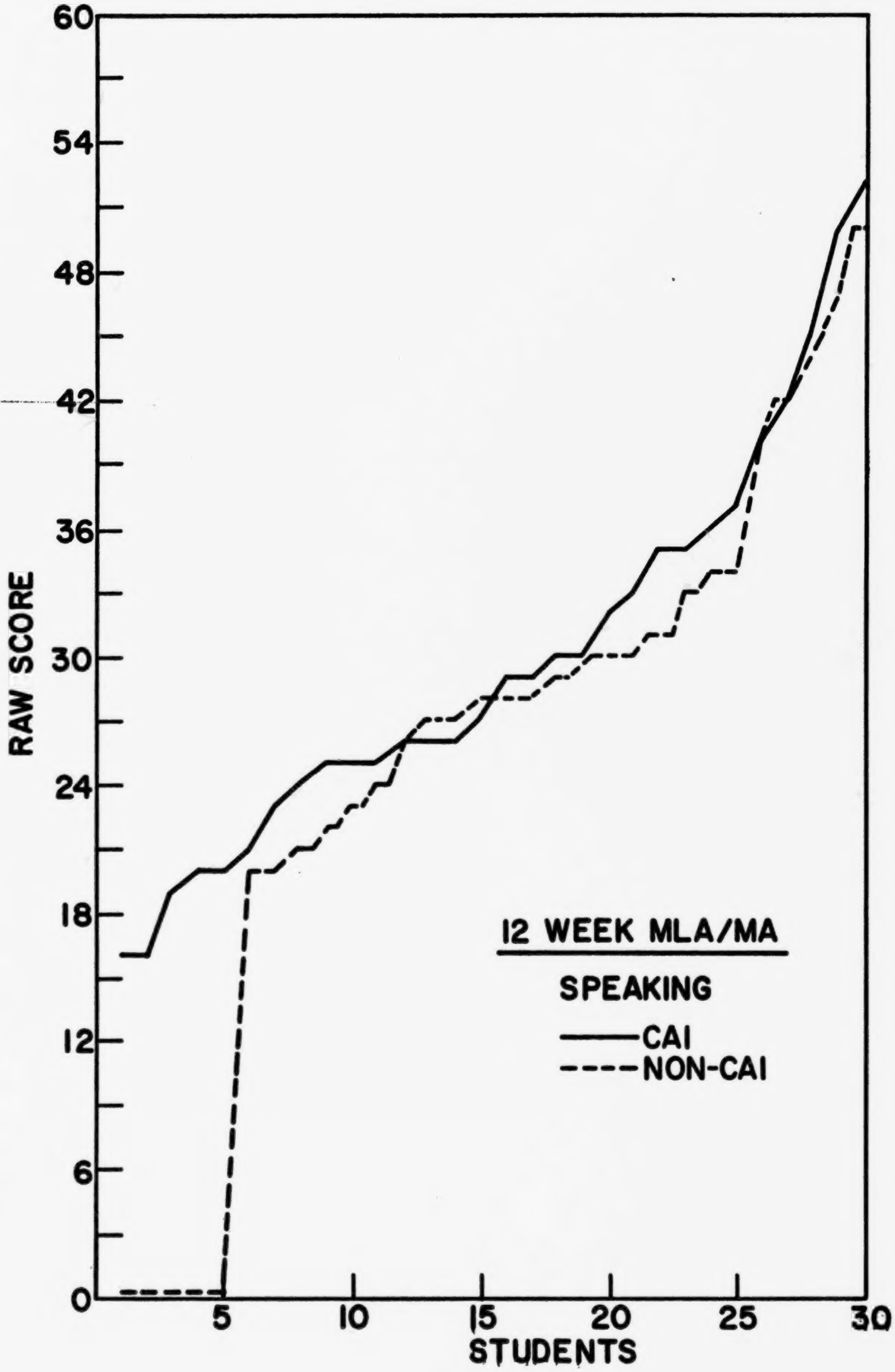


Figure III-C-10

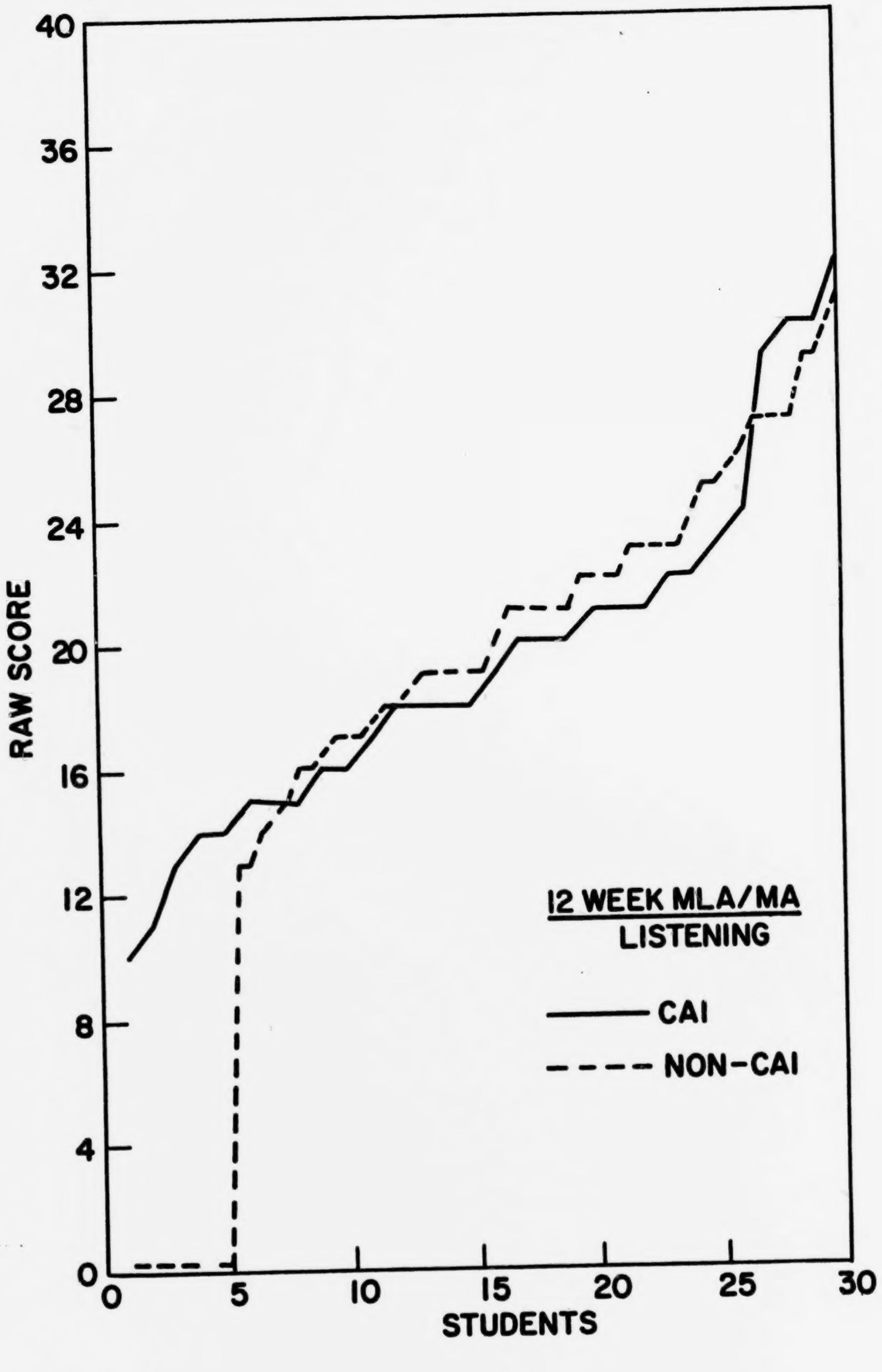


Figure III-C-11

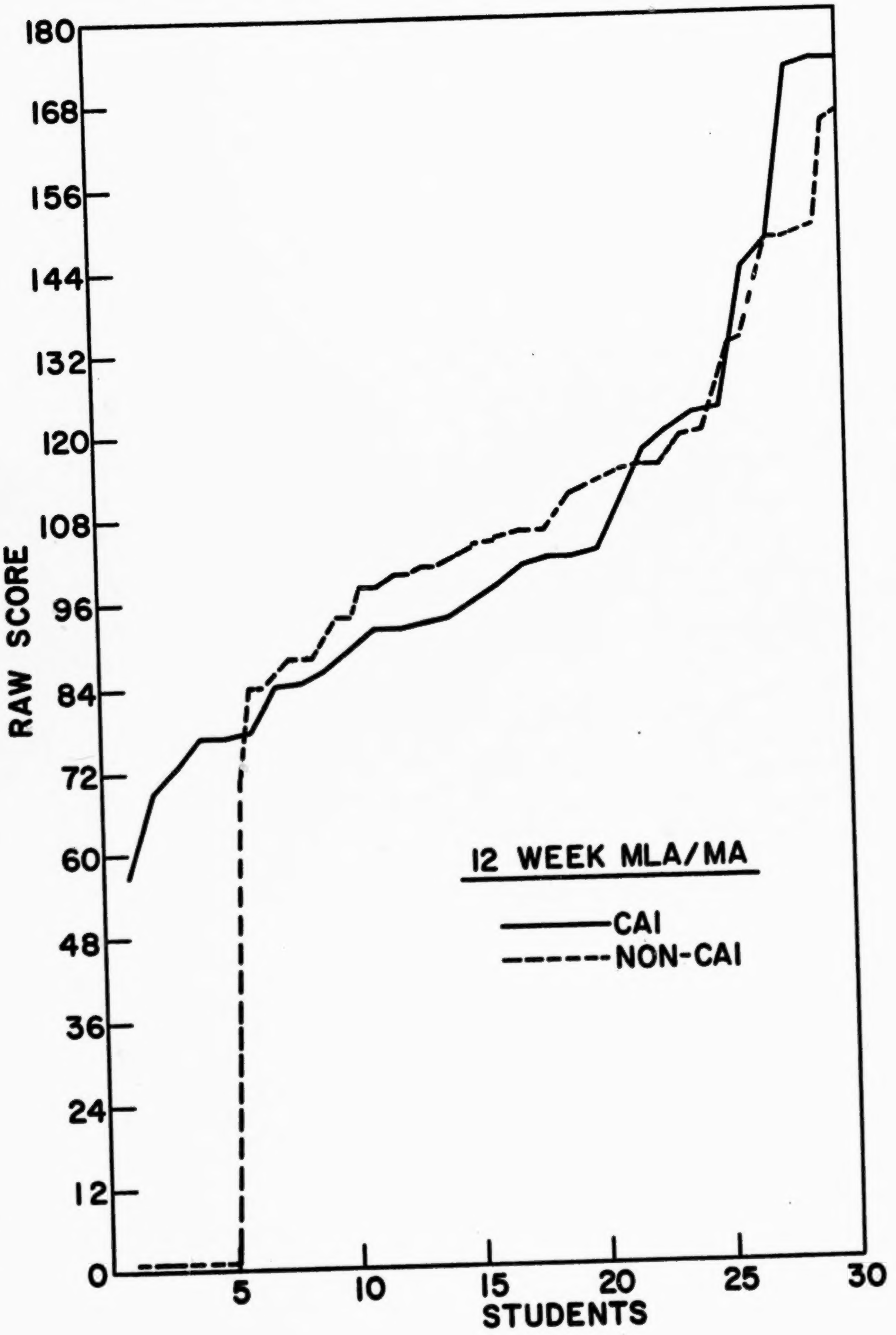


Figure III-C-12

Section D

This section contains performance comparisons for CAI and non-CAI sections on the tests and sub-tests administered after 18 weeks of instruction.

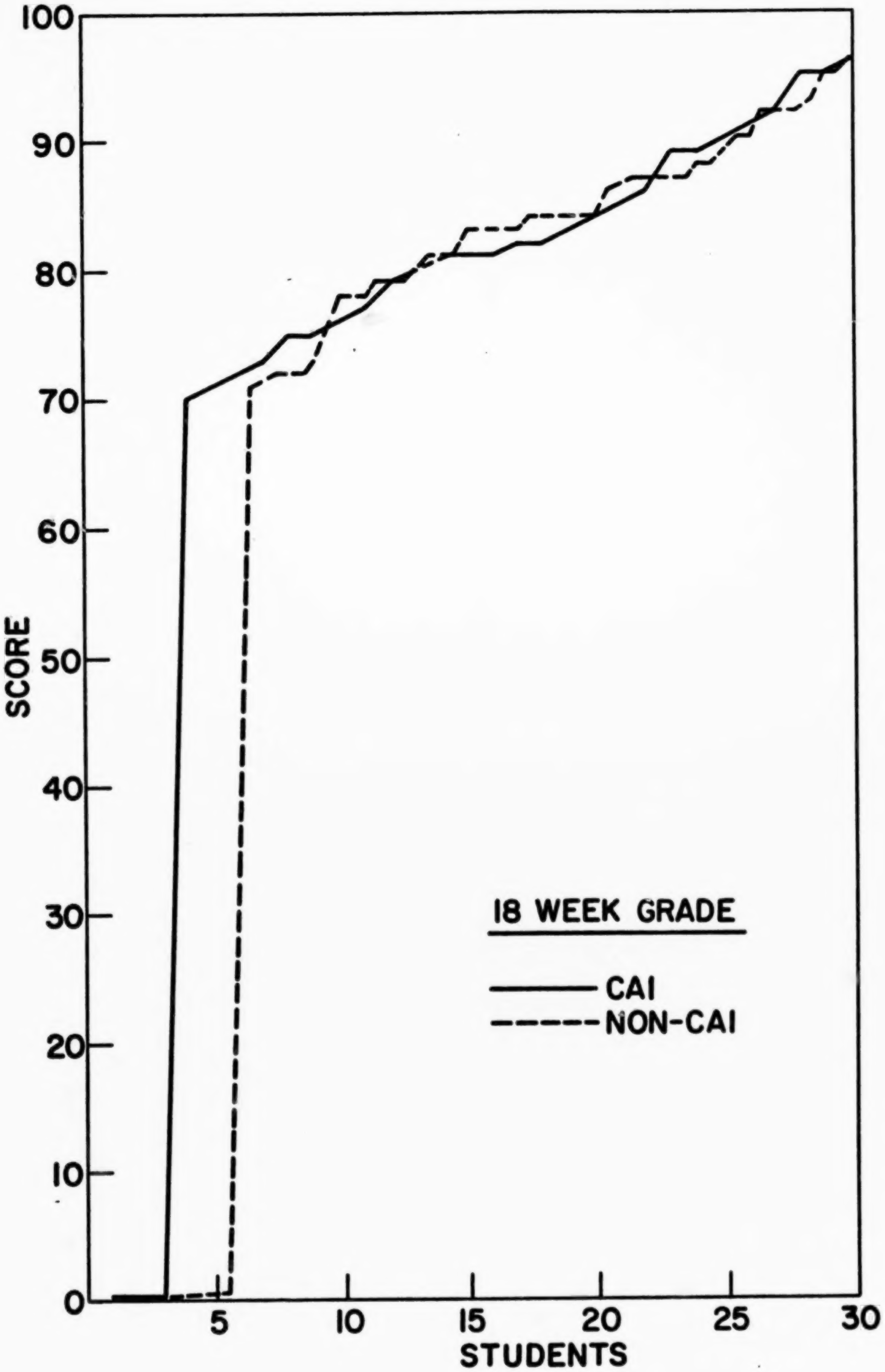


Figure III-D-1

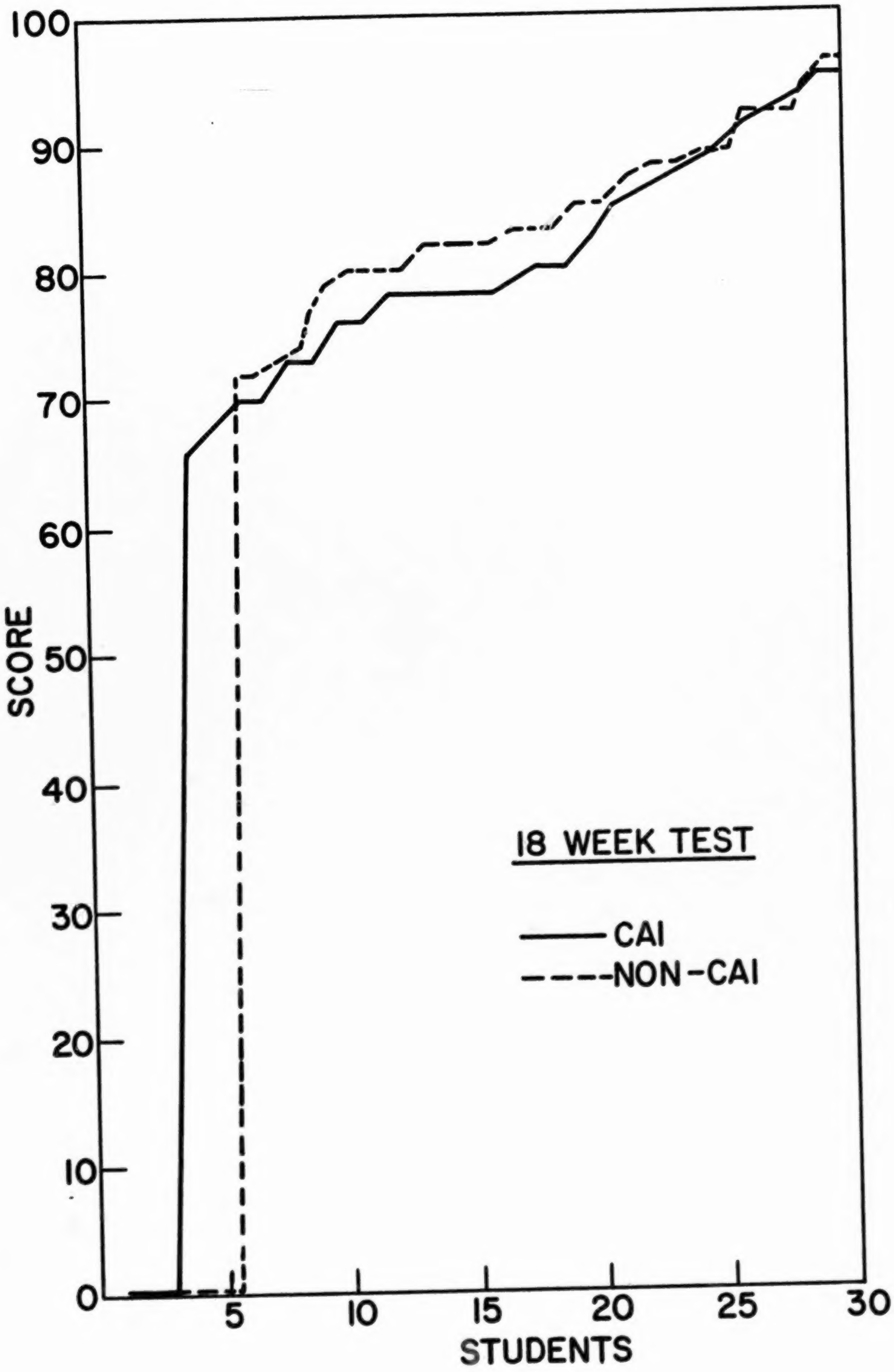


Figure III-D-2

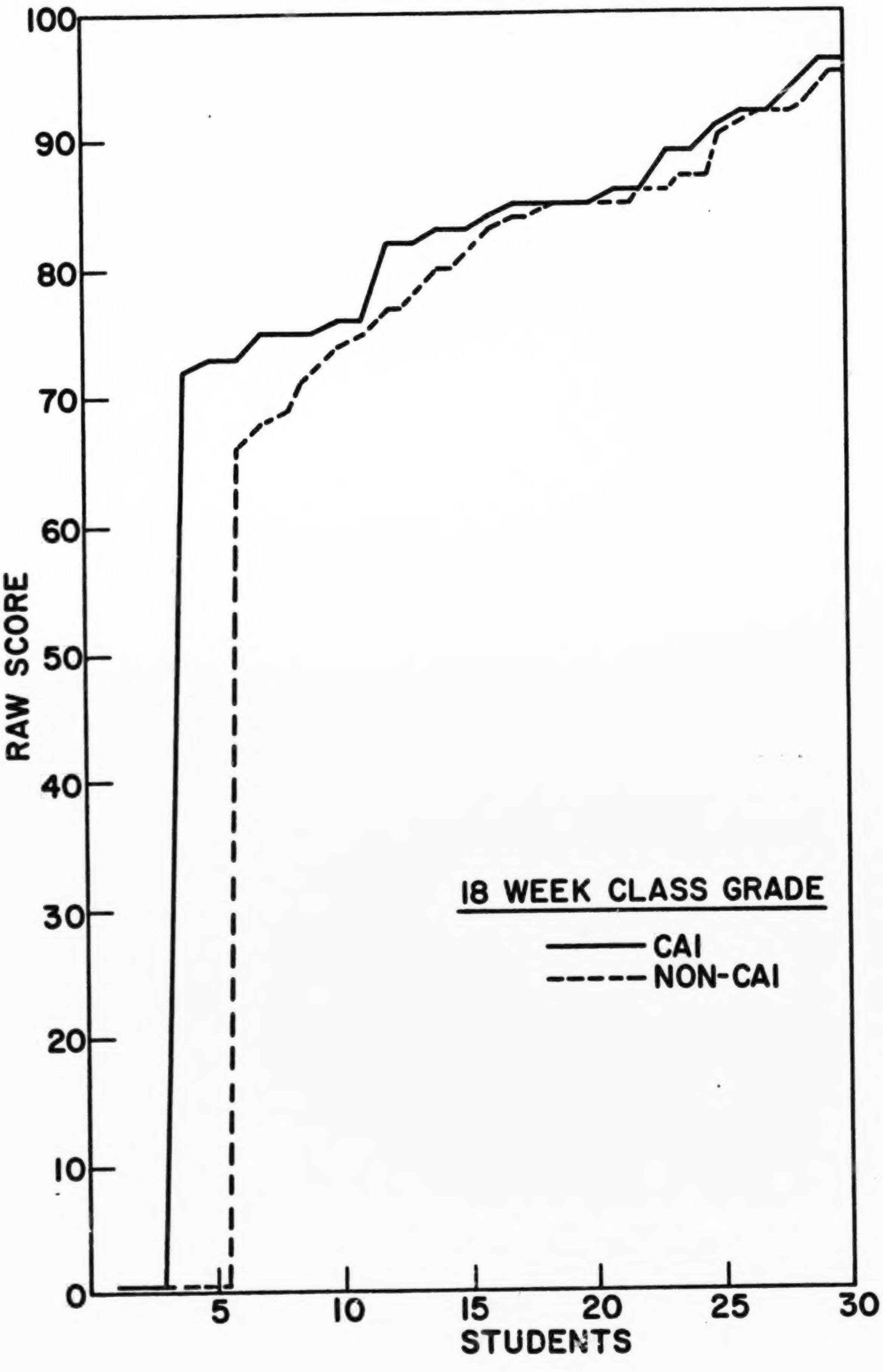


Figure III-D-3

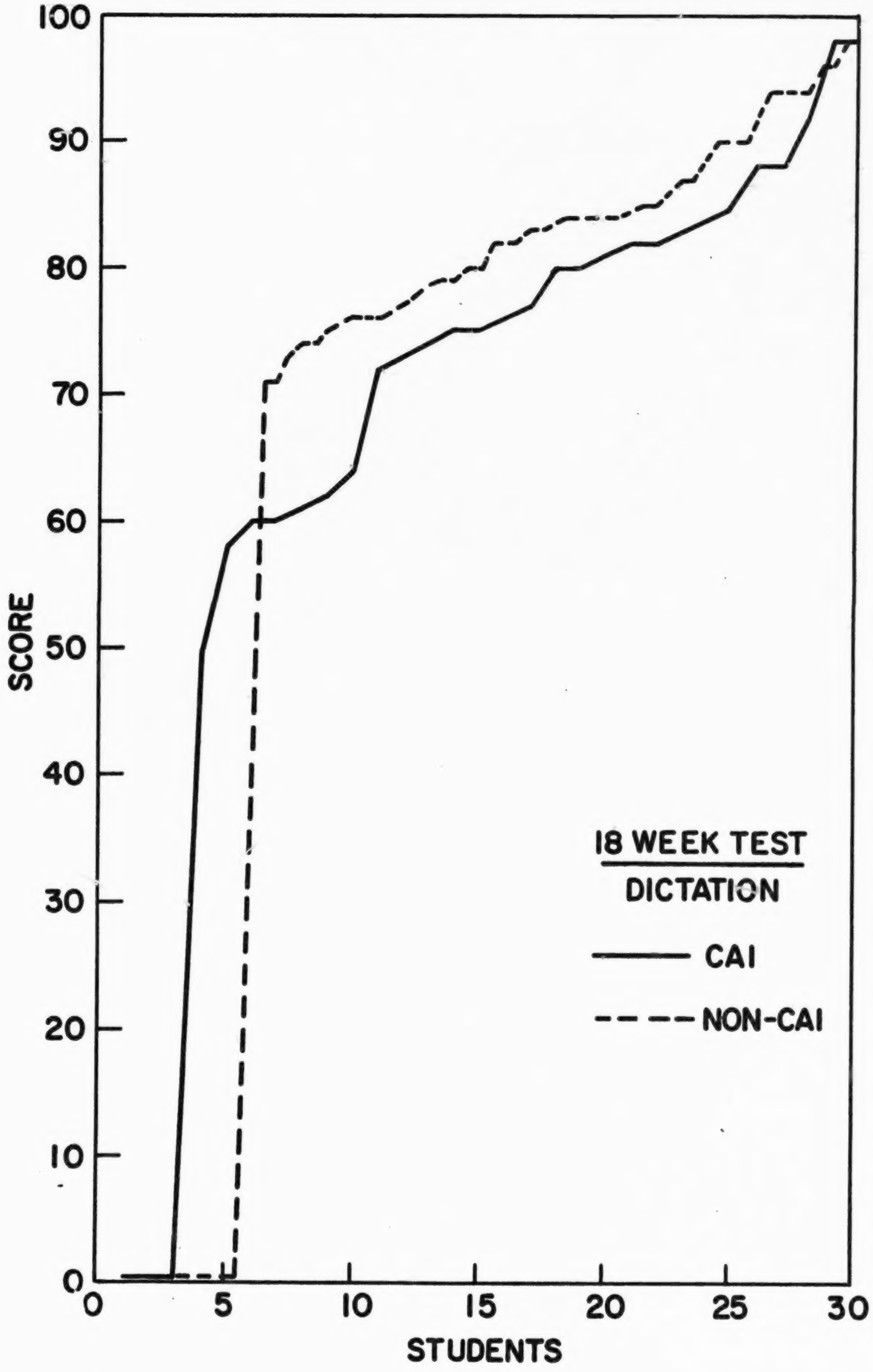


Figure III-D-4

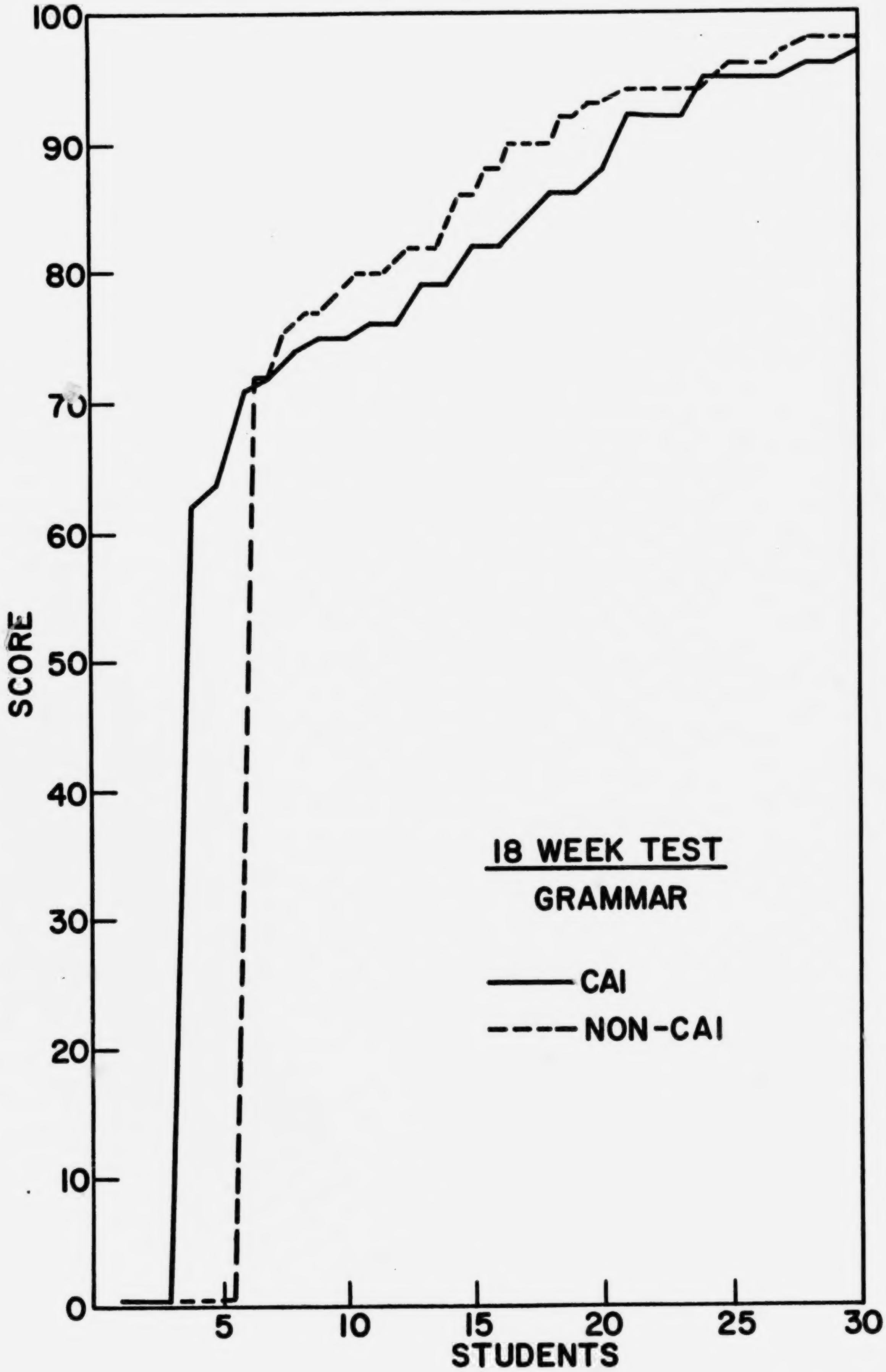


Figure III-D-5

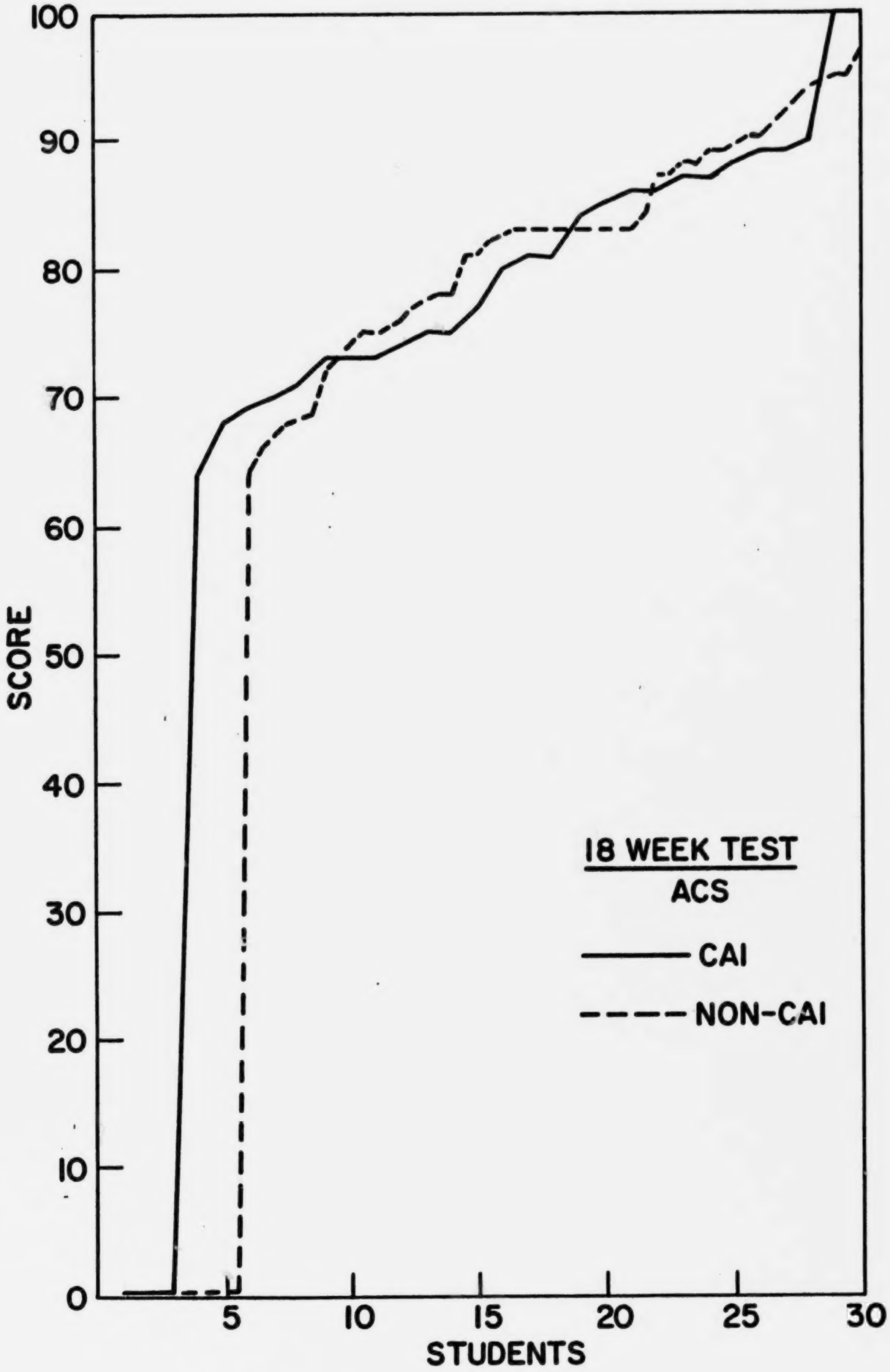


Figure III-D-6

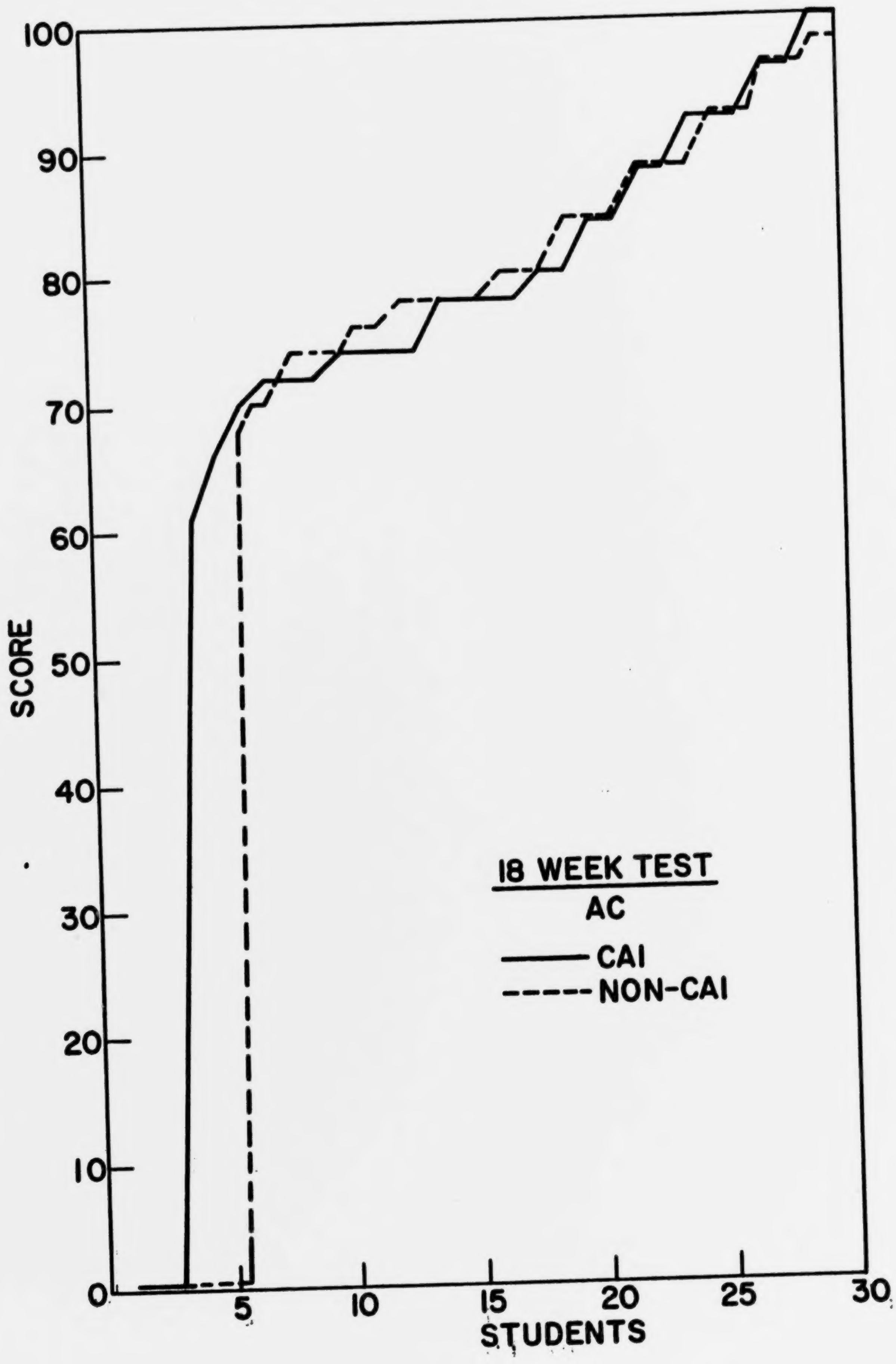


Figure III-D-7

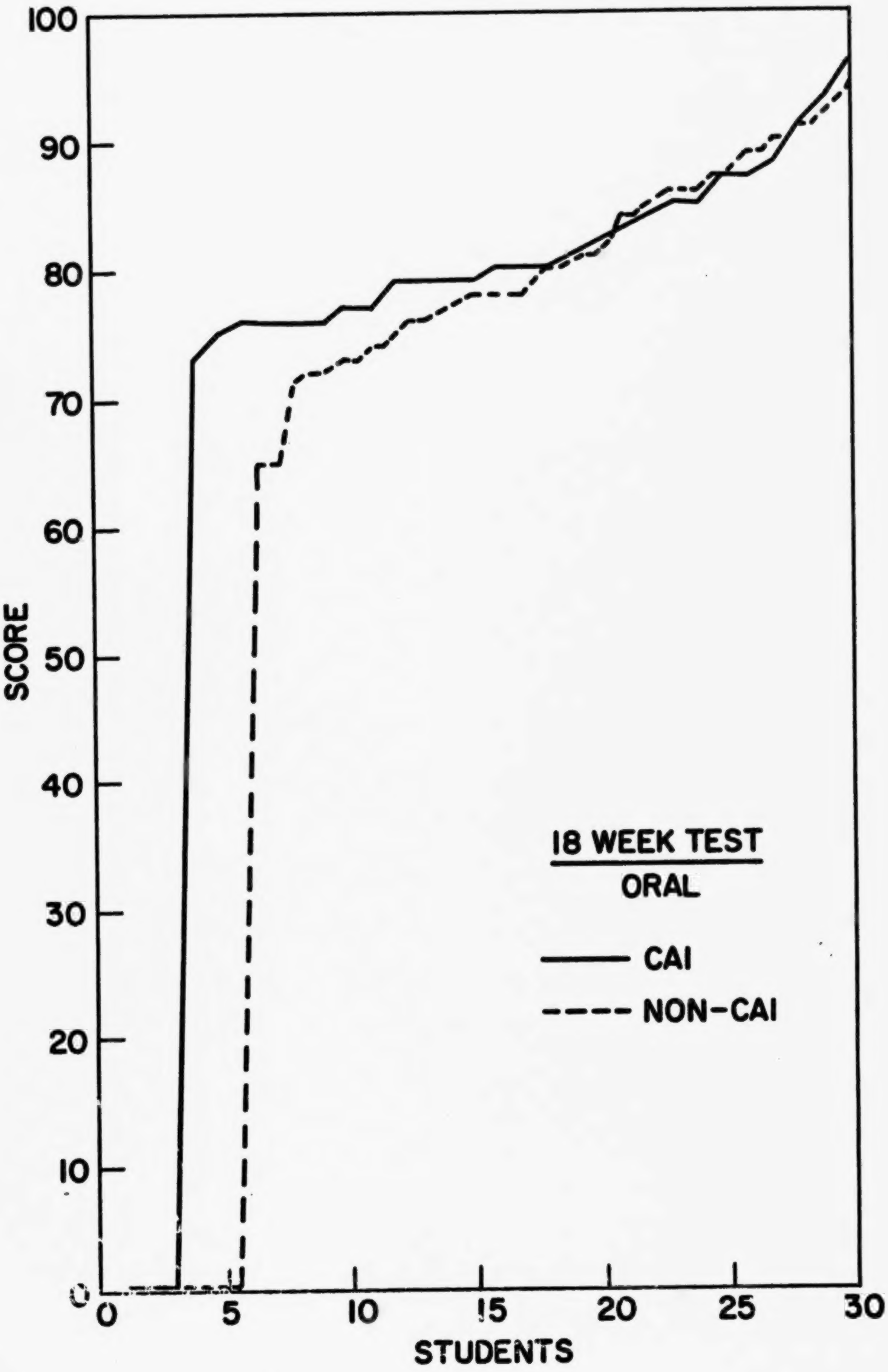


Figure III-D-8

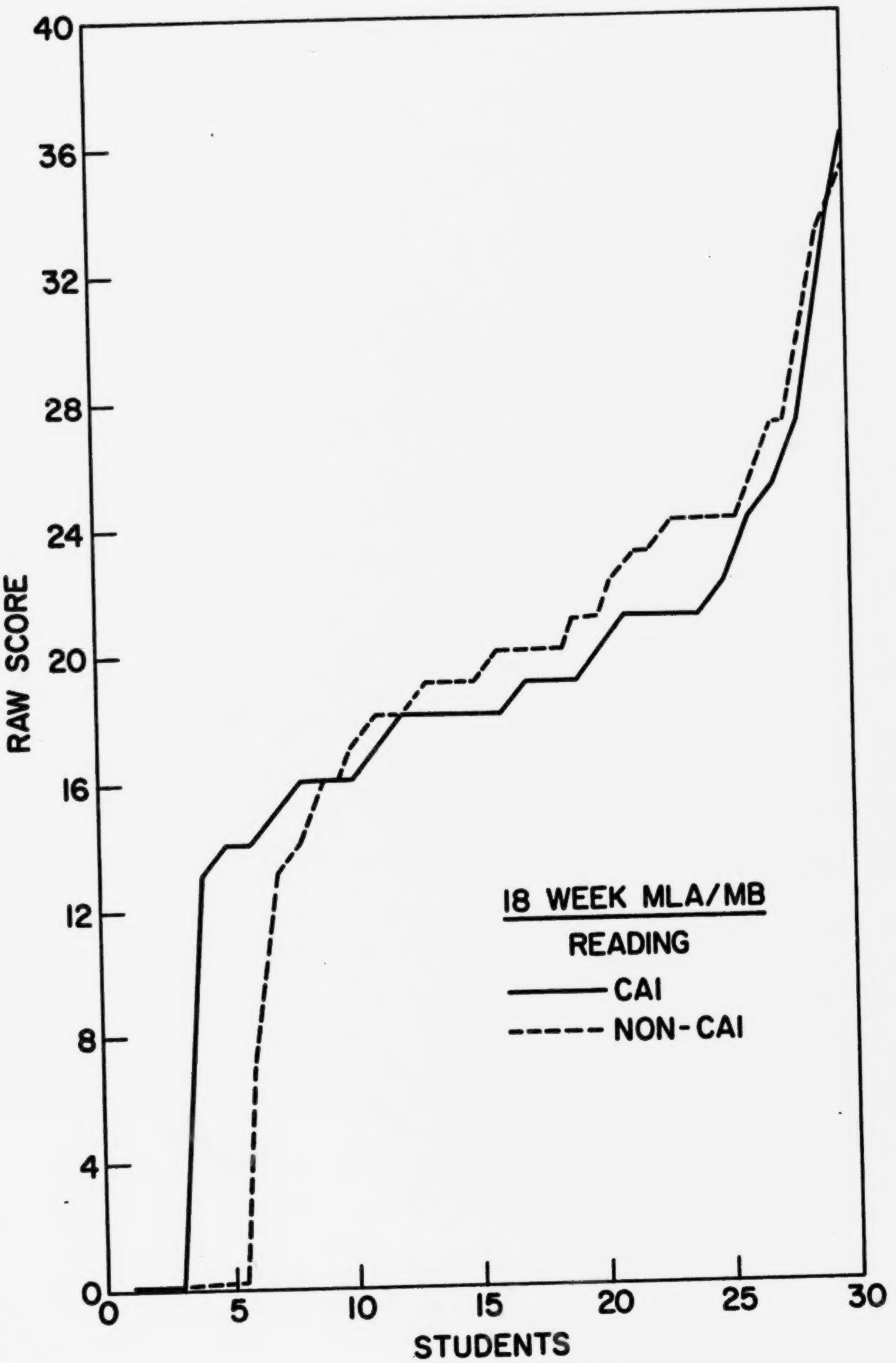


Figure III-D-9

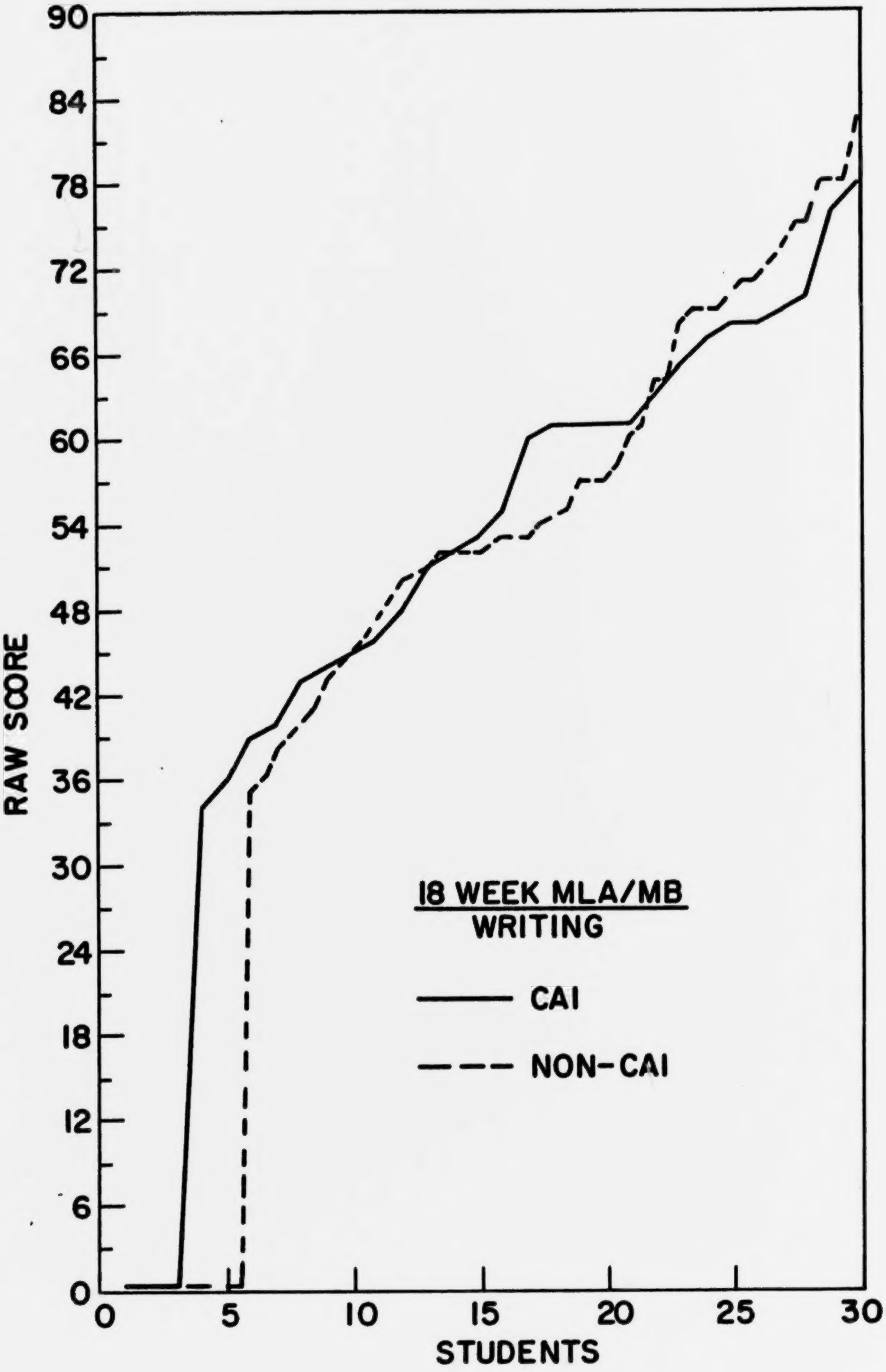


Figure III-D-10

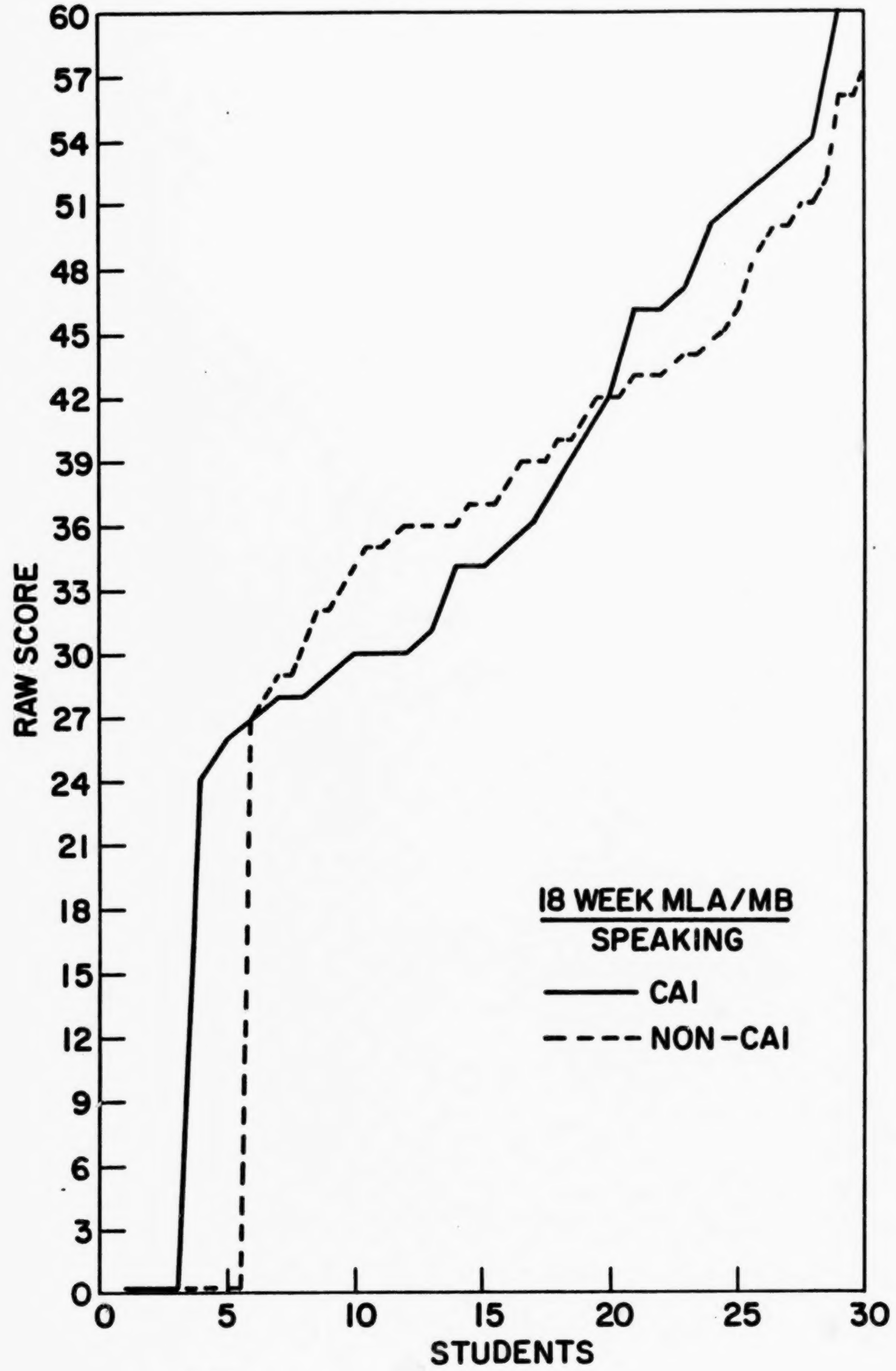


Figure III-D-11

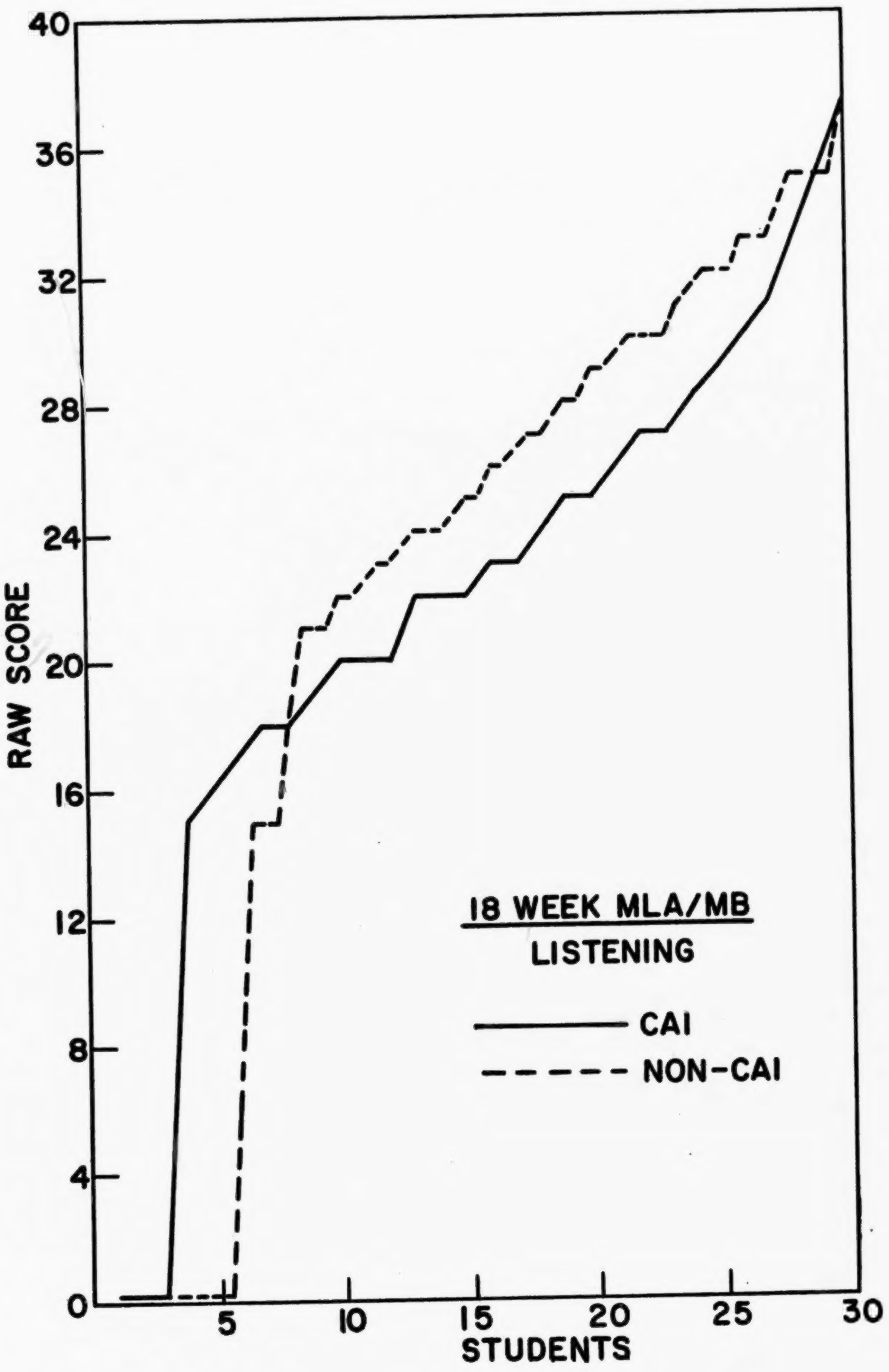


Figure III-D-12

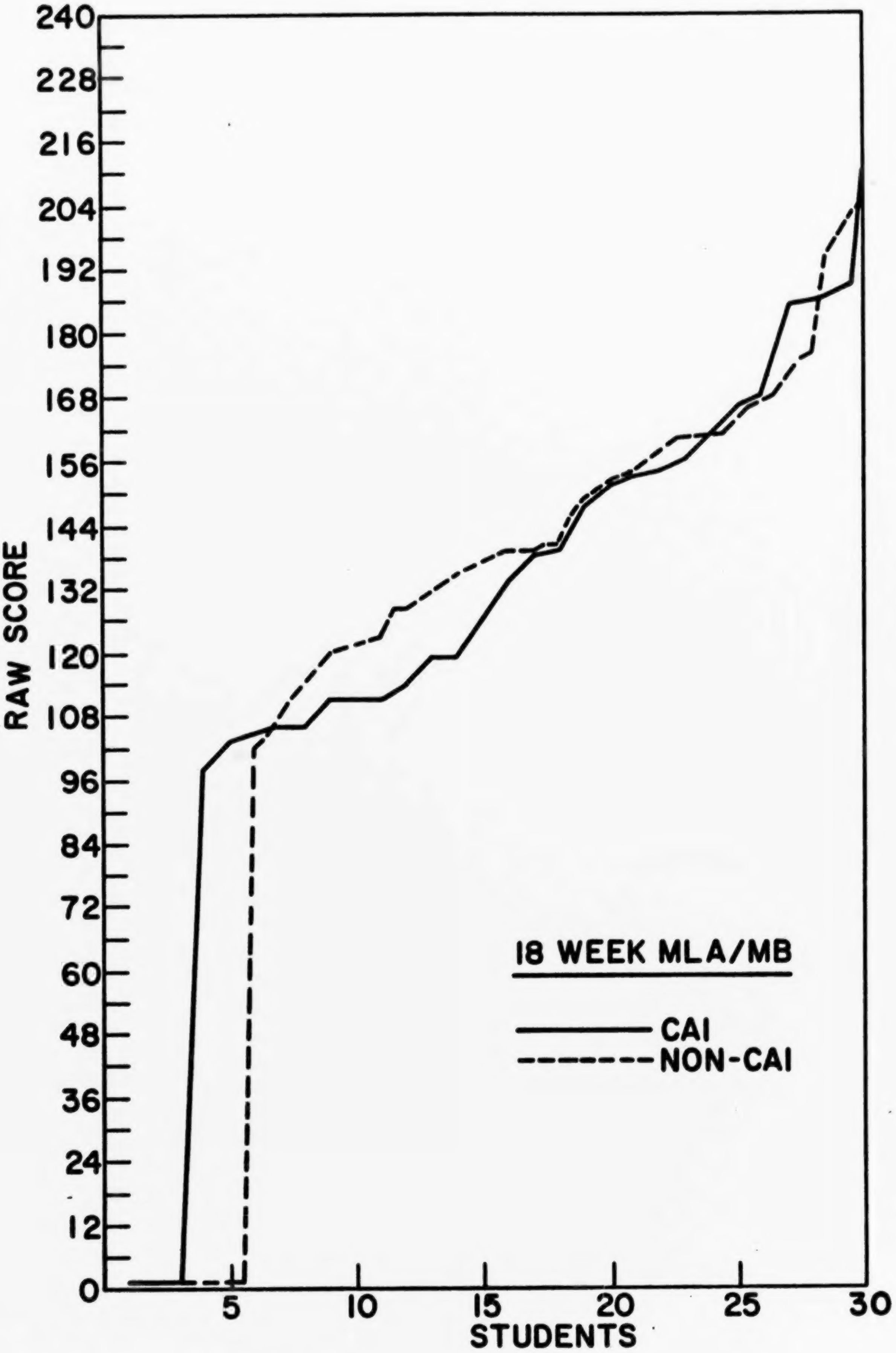
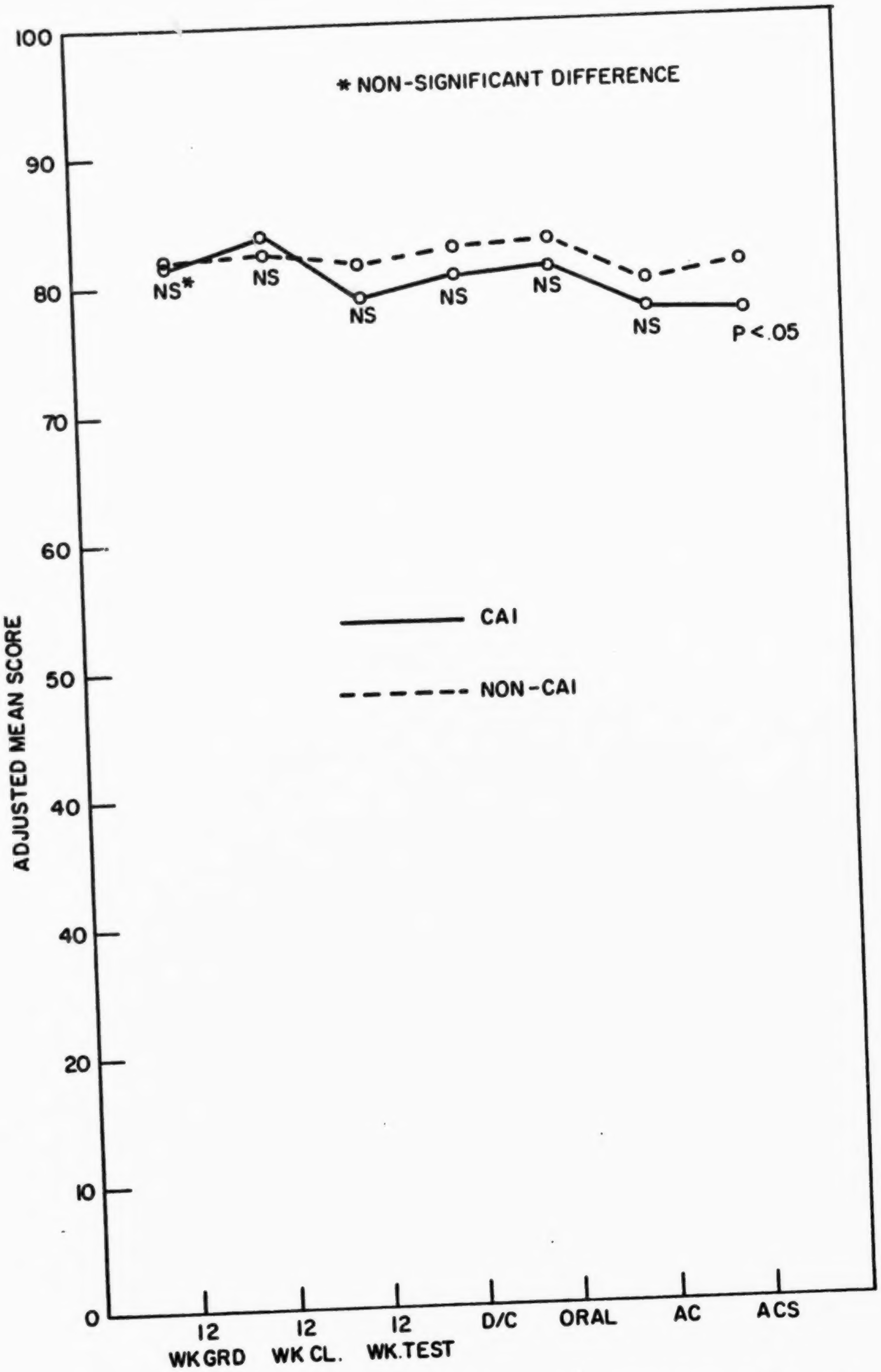


Figure III-D-13

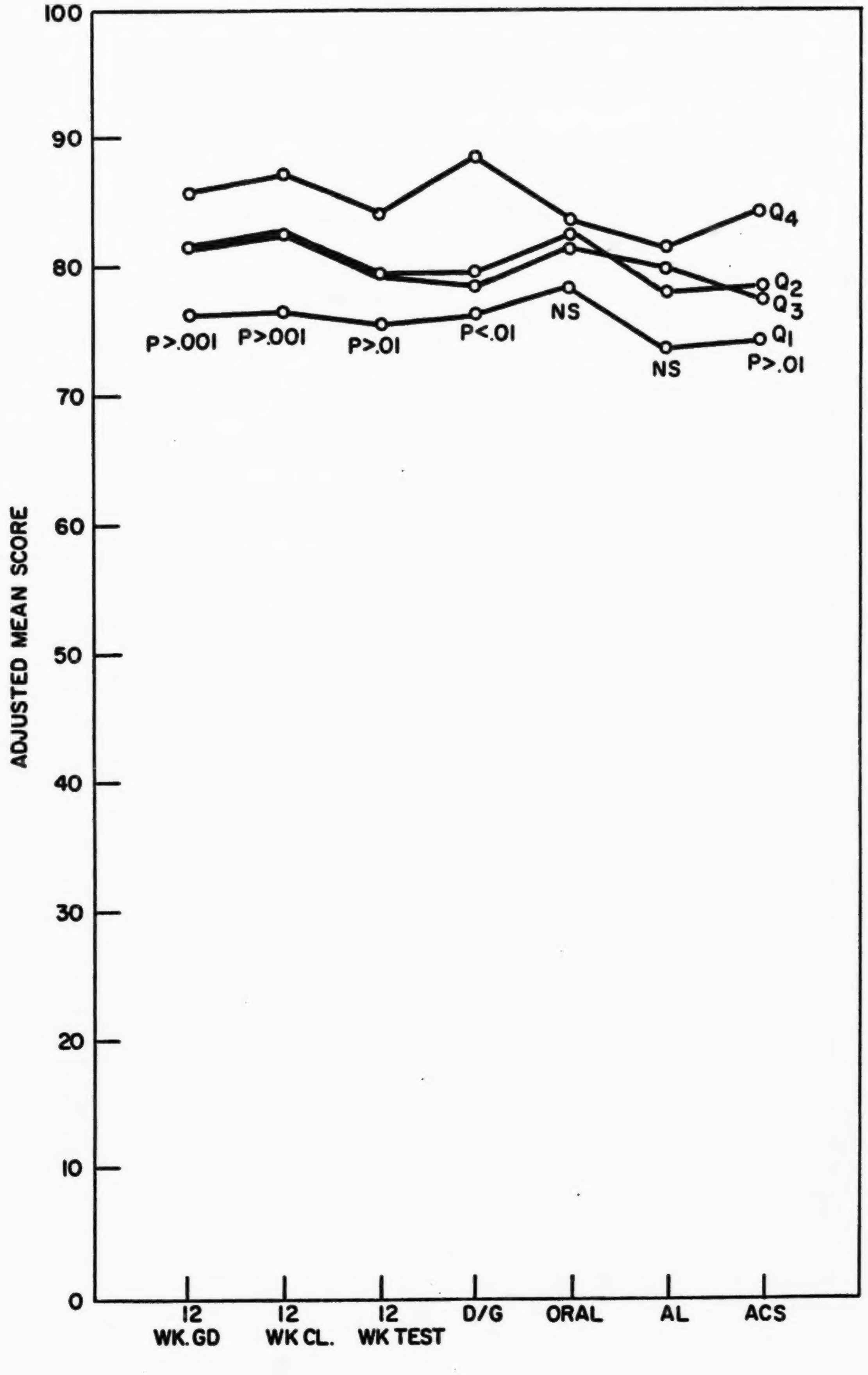
Section E

Contained in this section are the results of the principal analysis. Figs. III-E-1 through III-E-5 showed adjusted mean score comparisons on the 12 and 18 week DLI grades and the MLA Cooperative sub-tests Form MB, blocked by treatment (CAI vs. non-CAI) and by pooled quartiles. Figs. III-E-7 and III-E-8 contain tables of F-Test ratios blocked by quartiles. Figs. III-E-11 and III-E-12 present F-Test ratios based on a tercile blocking. Standard deviations, means, and magnitudes of the components of variance for the analysis by quartiles are given in Figs. III-E-9 and III-E-10.



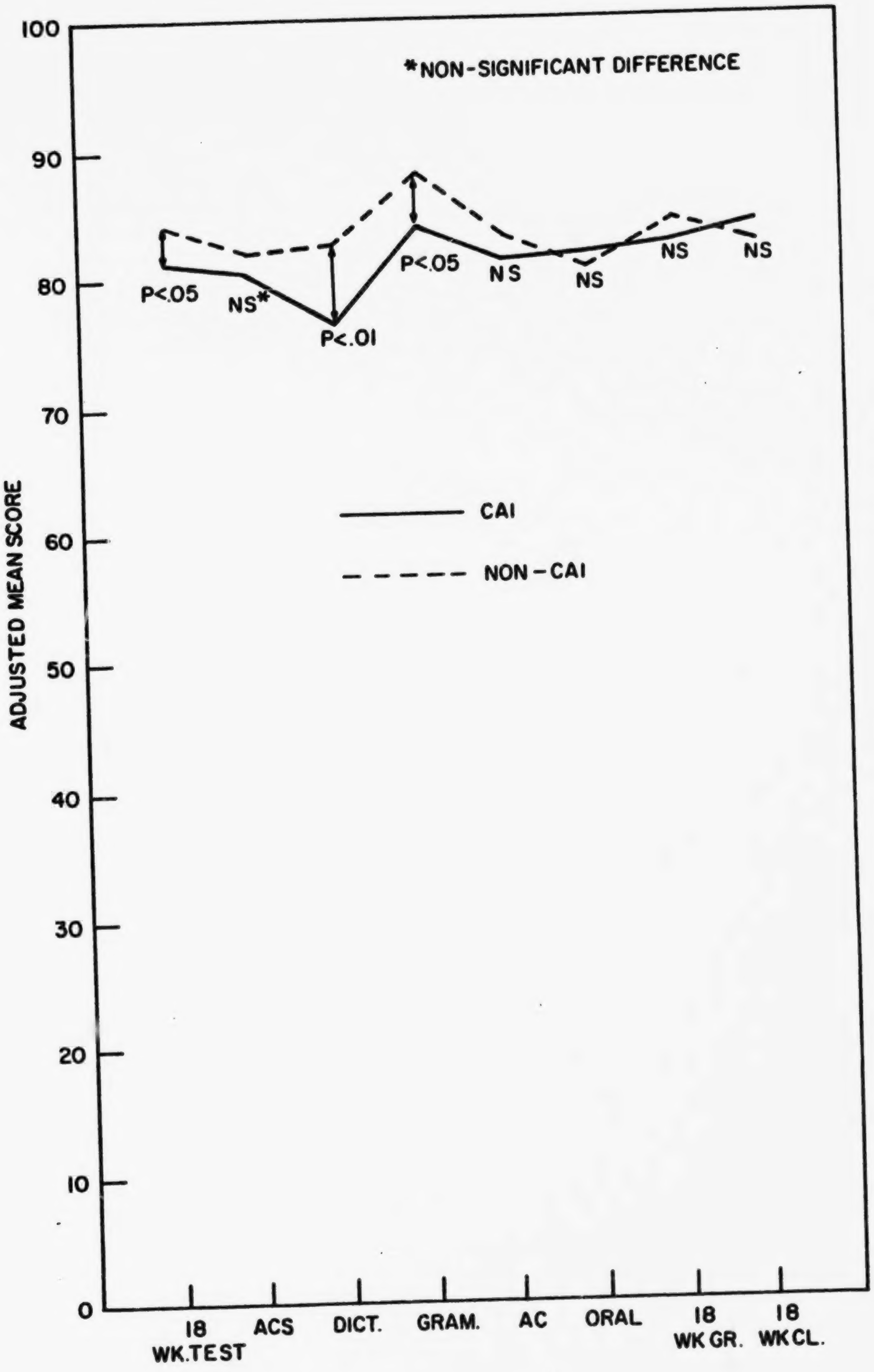
Adjusted Mean Score Comparisons for CAI and Non-CAI for the 12-Week DLI Grades.

Figure III-E-1



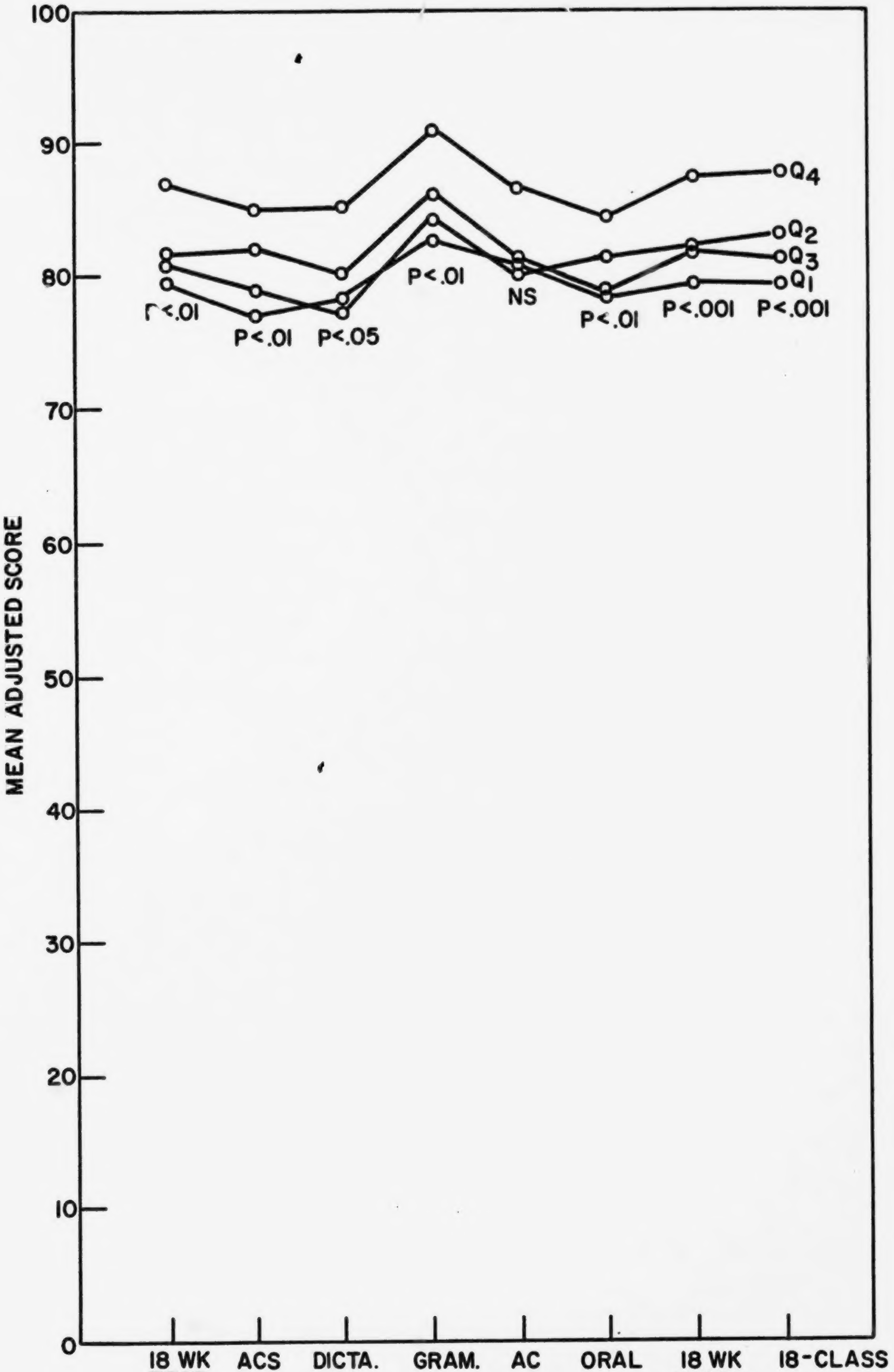
Adjusted Mean Performance of all Students for Each of the Four ALAT Quartiles for Each of the 12-Week DLI Grades.

Figure III-E-2



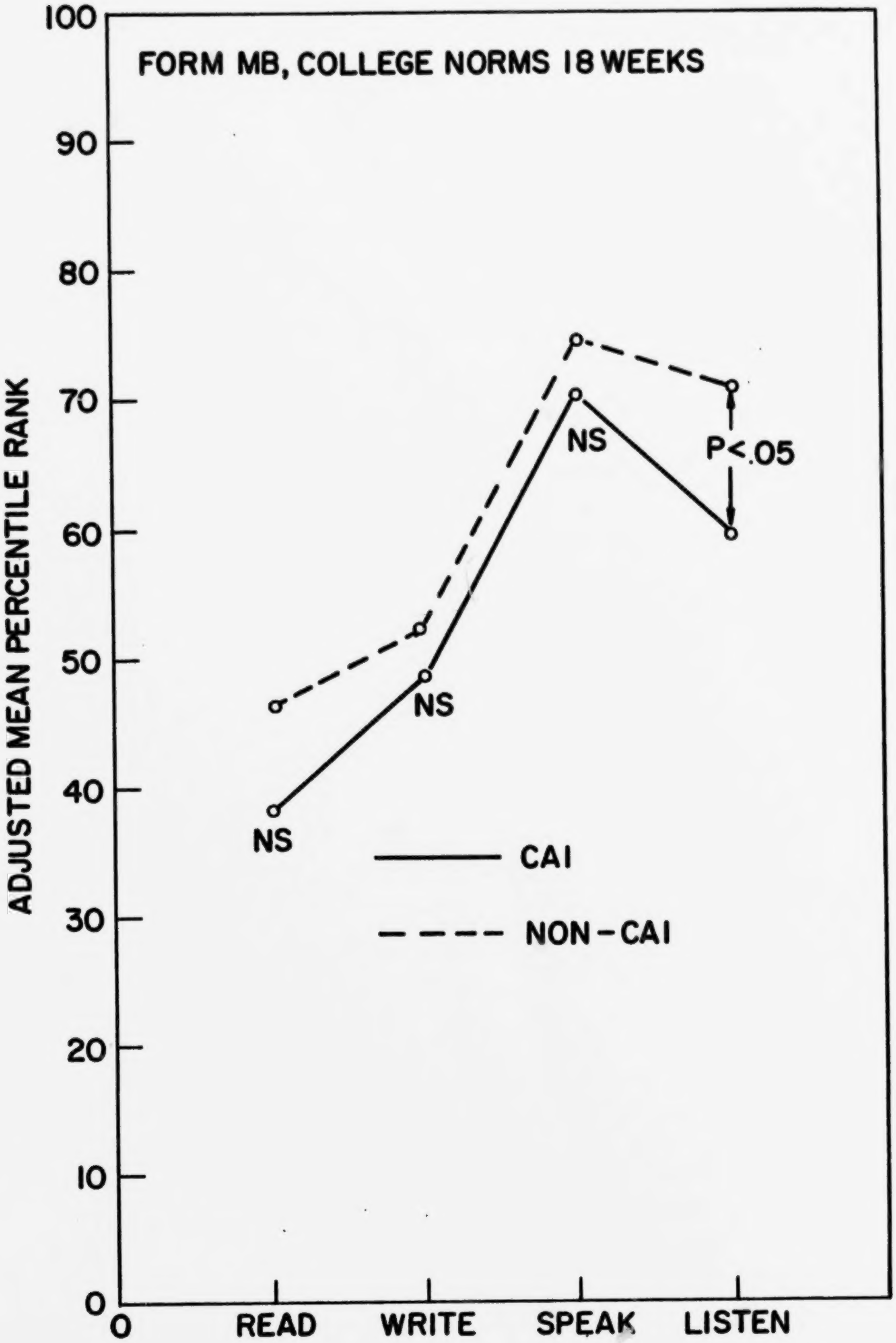
Adjusted Mean Score Comparisons for CAI and Non-CAI for the 18-Week DLI Grades

Figure III-E-3



Adjusted Mean Performance of All Students for Each of the Four ALAT Quartiles for Each of the 18-Week DLI Grades.

Figure III-E-4



Adjusted CAI and Non-CAI Group Means on MLA Cooperative Foreign Language Subtests.

Figure III-E-6

	DLI Grades	df	Total Grade	Class Grade	Total Test	Dictation/ Grammar	Dictation	Grammar	Oral	AC	ACS
12 Week	Treatment	1	0.7	0.4	3.9	1.0			2.3	1.3	5.5*
	ALAT	3	8.9‡	12.9‡	6.0†	8.5‡			2.0	1.7	7.1‡
	Tr x AL	3	1.1	0.9	1.3	0.6			0.8	1.1	3.1
	3 wk (res)	1	79.2‡	96.9‡	58.3‡	39.3‡			37.4‡	26.8‡	28.7‡
18 Week	Treatment	1	0.8	0.4	5.1*		10.4†	6.3*	0.2	0.0	0.7
	ALAT	3	6.4‡	6.3‡	6.0†		3.3*	4.3†	4.4†	1.7	4.2†
	Tr x AL	3	0.2	0.6	0.6		0.8	0.6	0.1	0.8	1.9
	3 wk (res)	1	44.5‡	45.7‡	36.1‡		20.3‡	15.3‡	35.8‡	14.6‡	18.4‡

Significance levels of F

For 1 df * $4.0 \leq F \leq 7.1$, $p \leq .05$
† $7.1 \leq F \leq 11.8$, $p \leq .01$
‡ $11.8 \leq F$ $p \leq .001$

For 3 df * $2.8 \leq F \leq 4.1$, $p \leq .05$
† $4.1 \leq F \leq 6.2$, $p \leq .01$
‡ $6.2 \leq F$ $p \leq .001$

Figure III-E-7 - F-Test Ratios for Analyses of Variance on 12 and 18 week DLI Grades Blocked by Quartiles

MLA Scores		df	Reading	Writing	Speaking	Listening	Total
12 Week Form A	Treatment	1	0.1	1.9	0.2	1.6	0.7
	ALAT	3	4.6†	19.0‡	5.4†	2.5	14.7‡
	Tr x AL	3	0.5	1.5	0.6	0.8	1.2
	3 wk (res)	1	0.5	30.7‡	14.0‡	18.6‡	27.5‡
18 Week Form B	Treatment	1	0.7	0.2	0.8	5.7*	1.6
	ALAT	3	4.1†	5.4†	7.0‡	3.4*	8.3‡
	Tr x AL	3	0.2	0.7	2.1	0.5	0.3
	3 wk (res)	1	3.1	14.9‡	15.8‡	12.7‡	19.8‡

Significance levels of F

For 1 df * $4.0 \leq F \leq 7.1$, $p \leq .05$
† $7.1 \leq F \leq 11.8$, $p \leq .01$
‡ $11.8 \leq F$ $p \leq .001$

For 3 df * $2.8 \leq F \leq 4.1$, $p \leq .05$
† $4.1 \leq F \leq 6.2$, $p \leq .01$
‡ $6.2 \leq F$ $p \leq .001$

Figure III-E-8 - F-Test Ratios for Analyses of Variance on 12 and 18 week MLA Scores Blocked by Quartiles

DLI Grades		Total Grade	Class Grade	Total Test	Dictation/ Grammar	Dictation	Grammar	Oral	AC	ACS
12 Week	Mean	81.7	82.9	80.3	81.2			82.1	78.9	78.8
	S. D.	8.9	8.3	9.8	11.5			9.9	13.3	10.6
	Treatment	0.6	0.4	1.6	1.0			1.4	1.5	2.2
	ALAT	3.4	3.6	3.3	5.0			2.2	2.9	4.2
	Tr x AL	1.2	1.0	1.6	1.4			1.4	2.4	2.8
	3 wk (res)	6.0	5.8	6.2	6.3			5.8	6.8	5.0
	Error	5.4	4.6	6.3	8.0			7.4	10.5	7.5
18 Week	Mean	83.5	83.4	83.2		81.0	86.6	81.4	82.6	81.8
	S. D.	7.0	7.4	7.4		10.4	9.1	6.8	9.4	8.6
	Treatment	.6	.5	1.6		3.4	2.5	.3	.2	.8
	ALAT	2.9	3.0	2.9		3.3	3.5	2.5	2.4	3.3
	Tr x AL	.5	.9	.9		1.6	1.2	.4	1.6	2.1
	3 wk (res)	4.5	4.8	4.3		4.8	3.9	4.2	4.1	4.0
	Error	5.2	5.4	5.5		8.3	7.6	5.4	8.4	7.3

Figure III-E-9 - Means, S. D. 's and Magnitudes of Components of Variance on 12 and 18 week DLI Grades Blocked by Quartiles

MLA Scores		Reading	Writing	Speaking	Listening	Total
12 Week Form A	Mean	19.4	39.1	30.4	20.6	109.4
	S.D.	5.4	12.1	9.0	5.0	26.5
	Treatment	.2	1.4	0.3	.7	1.9
	ALAT	2.3	7.5	3.8	1.5	14.7
	Tr x AL	.7	2.1	1.2	.8	4.2
	3 wk (res)	.4	5.6	3.4	2.3	11.9
	Error	5.0	8.1	7.3	4.3	18.0
18 Week Form B	Mean	20.9	51.9	40.5	26.0	144.8
	S.D.	5.6	13.2	9.0	5.9	28.0
	Treatment	.6	.6	.8	1.5	3.5
	ALAT	2.3	5.4	4.3	2.0	14.0
	Tr x AL	.5	1.9	2.3	.8	2.7
	3 wk (res)	1.2	5.4	3.8	2.3	12.7
	Error	5.2	10.9	7.4	5.0	22.0

Figure III-E-10 - Means, S.D.'s and Magnitudes of Components of Variance on 12 and 18 week MLA Test Scores Blocked by Quartiles

DLI Grades		df	Total Grade	Class Grade	Total Test	Dictation/ Grammar	Dictation	Grammar	Oral	AC	ACS
12 Week	Treatment	1	.0	1.5	1.8	.3			.9	.7	2.3
	ALAT	2	10.6‡	13.8‡	8.2‡	11.9‡			1.1	3.5*	9.6‡
	Tr x AL	2	1.0	.4	1.8	.4			.7	1.7	3.5*
	3 wk (res)	1	67.1‡	75.7‡	52.9‡	36.9‡			33.2‡	25.0‡	24.4‡
18 Week	Treatment	1	.4	1.1	4.5*		10.7†	4.1*	.7	.5	.3
	ALAT	2	10.5‡	8.2‡	11.5‡		6.5†	5.4†	6.1†	4.6*	6.7†
	Tr x AL	2	.3	.5	1.4		1.6	.6	.6	.5	2.6
	3 wk (res)	1	43.1‡	42.5‡	37.1‡		20.9‡	14.0‡	37.1‡	17.0‡	15.9‡

Significance levels of F

For 1 df	*	$4.0 \leq F \leq 7.1$,	$p \leq .05$
	†	$7.1 \leq F \leq 11.8$,	$p \leq .01$
	‡	$11.8 \leq F$	$p \leq .001$
For 2 df	*	$3.2 \leq F \leq 4.9$,	$p \leq .05$
	†	$5.0 \leq F \leq 7.8$,	$p \leq .01$
	‡	$7.9 \leq F$	$p \leq .001$

Figure III-E-11 - F-Test Ratios for Analysis of Variance on 12 and 18 week DLI Grades Blocked by Terciles

	MLA Scores	df	Reading	Writing	Speaking	Listening	Total
12 Week Form A	Treatment	1	.0	.1	.8	.8	.18
	ALAT	2	6.8†	18.9‡	7.2‡	2.9	17.0‡
	Tr x AL	2	.2	1.2	.8	.6	1.2
	3 wk (res)	1	.3	23.5‡	13.0‡	17.7‡	22.7‡
18 Week Form B	Treatment	1	1.2	.0	.1	4.5*	1.0
	ALAT	2	8.1‡	7.1†	12.0‡	4.8*	13.3‡
	Tr x AL	2	.1	.3	3.5*	.6	.3
	3 wk (res)	1	2.6	12.4‡	15.8‡	11.7†	18.0‡

Significance levels of F

For 1 df	*	$4.0 \leq F \leq 7.1$,	$p \leq .05$
	†	$7.1 \leq F \leq 11.8$,	$p \leq .01$
	‡	$11.8 \leq F$	$p \leq .001$
For 2 df	*	$3.2 \leq F \leq 4.9$,	$p \leq .05$
	†	$5.0 \leq F \leq 7.8$,	$p \leq .01$
	‡	$7.9 \leq F$	$p \leq .001$

Figure III-E-12 - F-Test Ratios for Analysis of Variance on 12 and 18 week MLA Scores Blocked by Terciles

Section F

This section provides a complete tabulation of means for the analysis of covariance. Presented for all criterion tests are original and adjusted (regression on 3-week test) means by quartiles for CAI, non-CAI, and pooled AC students.

	Original				Adjusted				
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Mean
Overall Grade									
CAI	75.5	78.5	78.3	88.6	73.2	80.9	81.8	87.0	81.5
Non-CAI	79.5	83.4	82.3	85.4	77.6	82.5	81.2	85.3	82.0
12 Week Test									
CAI	73.3	75.4	74.6	87.0	70.9	77.8	78.2	85.4	78.8
Non-CAI	79.5	83.3	81.7	83.9	77.6	82.4	80.6	83.8	81.3
Class Grade									
CAI	77.3	81.1	81.1	89.7	75.2	83.4	84.2	88.3	80.9
Non-CAI	79.0	82.9	82.4	86.9	77.1	82.0	81.2	86.9	82.4
ACS									
CAI	69.0	72.5	74.3	87.7	67.1	74.5	77.2	86.4	77.0
Non-CAI	79.0	84.3	78.6	83.6	77.4	83.5	77.7	83.6	80.7
AC									
CAI	74.3	72.5	72.5	86.6	71.6	75.2	76.5	84.8	77.3
Non-CAI	76.9	82.8	84.0	80.1	74.8	81.2	82.8	80.0	79.8
Dictation/Grammar									
CAI	74.0	77.3	73.8	90.7	71.6	79.7	77.5	89.0	80.2
Non-CAI	80.0	80.6	80.4	88.4	78.0	79.7	79.3	88.4	82.1
Oral									
CAI	75.5	79.3	78.4	84.9	73.3	81.5	81.7	83.3	81.0
Non-CAI	82.3	85.5	82.6	83.9	80.5	84.7	81.6	83.9	82.7
MLA-Reading									
CAI	17.0	17.3	18.1	23.3	16.8	17.4	18.4	23.2	19.0
Non-CAI	15.7	20.0	19.6	22.4	15.6	19.9	19.5	22.4	19.6
MLA-Writing									
CAI	26.3	35.5	30.6	52.1	24.1	37.7	34.0	50.6	37.9
Non-CAI	34.5	43.1	34.7	48.4	32.7	42.3	33.7	48.4	40.0

Figure III-F-1 - Adjusted Means by Quartiles for Analysis of Covariance on 12 Week Grades and Scores

Figure III-F-1 (Continued)

MLA-Speaking									
CAI	27.0	29.2	25.9	39.0	25.7	30.5	27.9	38.1	30.9
Non-CAI	28.3	31.9	27.4	33.5	27.2	31.4	26.8	33.5	29.9
MLA-Listening									
CAI	19.0	18.8	17.4	23.6	18.1	19.7	18.8	23.0	20.0
Non-CAI	19.8	23.3	20.6	22.2	19.1	22.9	20.2	22.2	21.1
MLA-Total									
CAI	85.3	100.3	92.0	138.0	84.0	100.7	92.8	137.7	105.1
Non-CAI	98.3	118.3	102.3	126.6	97.9	118.1	102.1	126.6	112.3

	Q1	Q2	Q3	Q4	Overall
Overall Grade	76.4	81.6	81.5	85.9	81.7
12 Week Test	75.7	79.7	79.6	84.3	80.2
Class Grade	76.5	82.8	82.8	87.3	82.9
ACS	74.5	78.3	77.5	84.5	79.2
AC	73.9	78.0	80.0	81.6	78.7
Dictation/Grammar	76.1	79.7	78.5	88.6	81.3
Oral	78.4	82.8	81.7	83.7	81.9
MLA Reading	15.9	18.5	19.0	22.7	19.3
MLA Writing	30.2	39.6	33.8	49.1	39.1
MLA Speaking	26.7	30.9	27.3	35.0	30.4
MLA Listening	18.8	21.1	19.6	22.5	20.7
MLA Total	94.2	108.0	97.9	130.3	109.3

Figure III-F-2 - Pooled Adjusted Means on
12 Week Grades and Scores

	Original				Adjusted				
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Mean
Overall Grade									
CAI	80.7	81.1	79.9	88.3	77.4	82.3	82.0	87.4	83.0
Non-CAI	81.9	83.1	83.1	87.4	80.8	82.9	82.7	87.8	84.0
18 Week Test									
CAI	79.3	78.3	78.4	87.7	76.2	79.4	80.5	86.9	81.3
Non-CAI	82.3	84.4	83.7	87.1	81.3	84.1	83.3	87.4	84.4
Class Grade									
CAI	81.7	83.5	80.4	88.3	78.2	84.7	82.7	87.4	84.2
Non-CAI	80.9	81.5	80.9	87.4	79.7	81.2	80.5	87.8	82.9
ACS									
CAI	77.7	75.7	79.9	89.4	74.7	76.8	81.8	88.7	80.9
Non-CAI	79.6	83.0	83.9	83.3	78.6	82.8	83.5	83.6	82.3
AC									
CAI	87.3	77.1	80.3	86.1	84.3	78.2	82.3	85.3	81.8
Non-CAI	81.1	83.4	81.2	87.3	80.1	83.1	80.8	87.7	83.4
Dictation									
CAI	75.3	72.2	74.0	85.3	71.8	73.5	76.3	84.4	76.8
Non-CAI	82.4	83.4	84.0	85.9	81.3	83.1	83.6	86.3	83.9
Grammar									
CAI	79.3	83.6	78.7	90.9	76.5	84.6	80.6	90.1	84.1
Non-CAI	86.2	88.6	87.3	91.2	85.3	88.4	87.0	91.6	88.4
Oral									
CAI	81.3	80.8	78.0	86.6	78.3	81.9	80.0	85.8	82.0
Non-CAI	80.1	81.5	79.1	84.0	79.1	81.2	78.7	84.4	81.2
MLA-Reading									
CAI	19.3	18.9	18.1	23.4	18.5	19.2	18.7	23.2	20.0
Non-CAI	17.9	20.3	21.1	25.4	17.6	20.2	21.0	25.4	21.6

Figure III-F-3 - Adjusted Means by Quartiles for Analysis of Covariance on 18 Week Grades and Scores

Figure III-F-3 (Continued)

	Original				Adjusted				
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Mean
MLA-Writing									
CAI	49.0	53.4	55.7	63.9	45.0	54.8	58.3	62.8	56.7
Non-CAI	53.7	53.4	55.0	65.9	52.4	53.1	54.5	66.3	57.8
MLA-Speaking									
CAI	35.0	34.6	36.6	50.6	32.2	35.6	38.4	49.8	39.6
Non-CAI	40.2	40.3	39.8	43.9	39.3	40.0	39.5	44.2	41.1
MLA-Listening									
CAI	22.7	23.3	21.3	28.9	21.0	23.9	22.4	28.4	24.4
Non-CAI	25.8	27.6	26.8	28.9	25.2	27.5	26.6	29.1	27.3
MLA-Total									
CAI	126.0	130.2	131.7	166.7	125.1	130.5	132.3	166.5	139.7
Non-CAI	137.6	141.5	142.7	164.0	137.2	141.4	142.6	164.1	148.5

	Q1	Q2	Q3	Q4	Overall
Overall Grade	79.9	82.5	82.4	87.6	83.6
18 Week Test	80.0	81.5	82.1	87.3	83.2
Class Grade	79.3	83.2	81.4	87.6	83.4
ACS	77.6	79.4	82.8	85.3	81.7
AC	81.1	80.4	81.4	86.9	82.8
Dictation	78.9	77.7	80.6	85.6	81.1
Grammar	83.1	86.3	84.3	91.1	86.7
Oral	78.9	81.6	79.2	84.8	81.5
MLA Reading	17.8	19.6	20.1	24.7	21.0
MLA Writing	50.5	54.0	56.1	65.2	57.4
MLA Speaking	37.5	37.6	39.0	46.1	40.5
MLA Listening	24.2	25.5	24.9	28.9	26.1
MLA Total	134.2	135.4	138.4	164.9	145.0

Figure III-F-4 - Pooled Adjusted Means on 18 Week Grades and Scores

Section G

This section presents matrices of correlations and intercorrelation for CAI and, where applicable, for non-CAI on the following measures:

1) 12 and 18 week DLI grades
2) 12 and 18 week MLA scores
3) CAI internal scores (i.e., 12 and 18 week recency scores, first attempt scores, time)
4) ALAT and 3-week test
5) Weekly time averages
6) Weekly time recency scores

	Grade (12)	Test (12)	Class (12)	AC (12)	ACS (12)	Dictation/ Grammar (12)	Oral (12)	Grade (18)	Test (18)	Class (18)	AC (18)	ACS (18)	Grammar (18)	Dictation (18)	Oral (18)
12 Week Grades															
Overall Grade	–	.97	.96	.85	.8	.91	.83	.97	.94	.95	.79	.71	.8	.82	.7
Test	.97	-	.86	.9	.87	.93	.74	.97	.93	.95	.78	.76	.75	.84	.69
Class Grade	.96	.86	-	.72	.64	.82	.87	.89	.87	.89	.73	.61	.78	.74	.64
Subtests															
AC	.85	.9	.72	-	.73	.79	.57	.86	.84	.82	.71	.81	.64	.68	.67
ACS	.8	.87	.64	.73	-	.72	.56	.81	.79	.8	.73	.61	.51	.79	.62
Dictation/Grammar	.91	.93	.82	.79	.72	-	.65	.9	.86	.9	.63	.68	.77	.78	.55
Oral	.83	.74	.87	.57	.56	.65	-	.8	.77	.78	.67	.46	.74	.68	.54
18 Week Grades															
Overall Grade	.97	.97	.89	.86	.81	.9	.8	-	.98	.97	.78	.77	.83	.89	.7
Test	.94	.93	.87	.84	.79	.86	.77	.98	-	.91	.8	.79	.83	.91	.72
Class Grade	.95	.95	.89	.82	.8	.9	.78	.97	.91	-	.72	.69	.8	.81	.62
Subtests															
AC	.79	.78	.73	.71	.73	.63	.67	.78	.8	.72	-	.55	.52	.7	.61
ACS	.71	.76	.61	.81	.61	.68	.46	.77	.79	.69	.55	-	.5	.7	.51
Grammar	.8	.75	.78	.64	.51	.77	.74	.83	.83	.8	.52	.5	-	.62	.62
Dictation	.82	.84	.74	.68	.79	.78	.68	.89	.91	.81	.7	.7	.62	-	.52
Oral	.7	.69	.64	.67	.62	.55	.54	.7	.72	.62	.61	.51	.62	.52	-

Figure III-G-1 - Intercorrelations of 12 and 18 Week DLI Grades for CAI

		READ (12)	WRITE (12)	SPEAK (12)	LISTEN (12)	TOTAL (12)	READ (18)	WRITE (18)	SPEAK (18)	LISTEN (18)	TOTAL (18)
READ	(12)	-	.55	.58	.57	.59	.42	.57	.65	.72	.59
WRITE	(12)	.55	-	.83	.69	.76	.81	.75	.77	.94	.87
SPEAK	(12)	.58	.83	-	.7	.79	.74	.89	.83	.93	.9
LISTEN	(12)	.57	.69	.7	-	.54	.59	.67	.76	.82	.71
TOTAL	(12)	.59	.76	.79	.54	-	.58	.68	.69	.8	.77
READ	(18)	.42	.81	.74	.59	.58	-	.76	.78	.78	.91
WRITE	(18)	.57	.75	.89	.67	.68	.76	-	.84	.85	.93
SPEAK	(18)	.65	.77	.83	.76	.69	.78	.84	-	.87	.91
LISTEN	(18)	.72	.94	.93	.82	.8	.78	.85	.87	-	.91
TOTAL	(18)	.59	.87	.9	.71	.77	.91	.93	.91	.91	-

Figure III-G-2 - Intercorrelations MLA Forms MA and MB for CAI

	Av. Time (12)	Grammar (12)	Translation(12)	Dictation (12)	AC (12)	Total (12)	Av. Time (18)	Grammar (18)	Translation (18)	Dictation (18)	AC (18)	Total (18)	Grammar (1st attempt)	Translation (1st attempt)	Dictation (1st attempt)	AC (1st attempt)	Total (1st attempt)
12 Week Internal Variables																	
Av. Time	-	.7	.77	.76	.76	.81	.95	.62	.73	.61	.65	.74	.63	.74	.6	.72	.72
Grammar	.7	-	.85	.81	.79	.93	.68	.77	.7	.77	.58	.79	.9	.89	.84	.89	.93
Translation	.77	.85	-	.71	.84	.91	.75	.79	.78	.68	.66	.82	.83	.93	.74	.84	.89
Dictation	.76	.81	.71	-	.84	.91	.76	.65	.83	.8	.71	.84	.69	.73	.76	.82	.8
AC	.76	.79	.84	.84	-	.94	.78	.79	.83	.7	.78	.87	.77	.84	.71	.88	.87
Total	.81	.93	.91	.91	.94	-	.81	.81	.85	.8	.74	.9	.85	.91	.82	.93	.94
18 Week Internal Variables																	
Av. Time	.95	.68	.75	.76	.78	.81	-	.63	.7	.68	.67	.75	.59	.72	.64	.74	.73
Grammar	.62	.77	.79	.65	.79	.81	.63	-	.73	.77	.61	.88	.87	.87	.76	.79	.87
Translation	.73	.7	.78	.83	.83	.85	.7	.73	-	.71	.83	.93	.72	.82	.71	.78	.81
Dictation	.61	.77	.68	.8	.7	.8	.68	.77	.71	-	.66	.87	.66	.73	.9	.79	.81
AC	.65	.58	.66	.71	.78	.74	.67	.61	.83	.66	-	.87	.63	.69	.61	.77	.73
Total	.74	.79	.82	.84	.87	.9	.75	.88	.93	.87	.87	-	.82	.88	.83	.88	.91
First Att. Variables																	
Grammar	.63	.9	.83	.69	.77	.85	.59	.87	.72	.66	.63	.82	-	.93	.74	.86	.94
Translation	.74	.89	.93	.73	.84	.91	.72	.87	.82	.73	.69	.88	.93	-	.81	.88	.97
Dictation	.6	.84	.74	.76	.71	.82	.64	.76	.71	.9	.61	.83	.74	.81	-	.81	.87
AC	.72	.89	.84	.82	.88	.93	.74	.79	.78	.79	.77	.88	.86	.88	.81	-	.96
Total	.72	.93	.89	.8	.87	.94	.73	.87	.81	.81	.73	.91	.94	.97	.87	.96	-

Matrix III-G-3 - Intercorrelation of Internal Scores

	Grammar (12)	Translation (12)	Dictation (12)	AC (12)	Total (12)	Time (12)	Grammar (18)	Translation (18)	Dictation (18)	AC (18)	Total (18)	Time (18)	Grammar (1st attempt)	Translation(1st attempt)	Dictation (1st attempt)	AC (1st attempt)	Total (1st attempt)
12 Week																	
Overall Grade	.76	.76	.86	.77	.86	.9	.92	.74	.73	.86	.78	.85	.64	.72	.71	.78	.77
Test	.74	.73	.86	.78	.85	.91	.93	.76	.76	.86	.77	.86	.65	.73	.72	.79	.78
Class Grade	.66	.64	.81	.73	.77	.86	.84	.66	.64	.81	.73	.77	.57	.64	.64	.69	.68
ACS	.65	.68	.68	.6	.71	.77	.88	.66	.69	.75	.73	.77	.57	.67	.59	.69	.68
AC	.66	.69	.75	.73	.77	.81	.78	.65	.68	.68	.6	.71	.59	.64	.51	.72	.68
Dictation/Grammar	.76	.7	.85	.71	.82	.81	.82	.76	.7	.85	.71	.82	.62	.67	.78	.75	.75
Oral	.49	.51	.71	.68	.65	.73	.73	.49	.51	.71	.68	.65	.44	.53	.57	.55	.56
18 Week																	
Overall Grade	.76	.74	.88	.75	.85	.88	.61	.81	.75	.74	.82	.9	.65	.74	.74	.78	.78
Test	.71	.67	.86	.7	.8	.85	.52	.75	.71	.72	.76	.88	.57	.66	.69	.73	.71
Class Grade	.79	.8	.86	.79	.88	.87	.69	.83	.77	.72	.85	.88	.73	.8	.76	.82	.83
ACS	.68	.65	.8	.76	.79	.72	.35	.47	.4	.48	.48	.71	.45	.47	.39	.48	.49
AC	.57	.49	.62	.52	.6	.69	.5	.65	.64	.59	.67	.78	.52	.62	.58	.75	.68
Dictation	.61	.52	.81	.54	.67	.71	.36	.67	.57	.6	.62	.72	.45	.51	.57	.61	.58
Grammar	.62	.63	.7	.62	.7	.72	.57	.73	.77	.72	.78	.72	.54	.66	.77	.65	.69
Oral	.41	.47	.47	.46	.49	.68	.31	.44	.48	.54	.49	.72	.32	.42	.38	.53	.46

Figure III-G-4 - Correlations between 12 and 18 Week DLI Grades and Internal Scores for CAI

	12 Week Time	12 Week Read	12 Week Write	12 Week Speaking	12 Week Listening	12 Week Total	18 Week Time	18 Week Read	18 Week Write	18 Week Speaking	18 Week Listening	18 Week Total
Internal Scores (12)												
Grammar	.7	.39	.82	.62	.67	.76	.68	.61	.75	.62	.61	.74
Translation	.77	.44	.87	.65	.61	.79	.75	.61	.78	.57	.72	.75
Dictation	.76	.41	.75	.7	.69	.76	.76	.54	.85	.77	.72	.84
AC	.76	.41	.74	.61	.62	.72	.78	.62	.76	.63	.7	.76
Total	.81	.44	.86	.7	.7	.82	.81	.64	.85	.7	.75	.84
Internal Scores (18)												
Grammar	.62	.39	.69	.54	.4	.62	.63	.61	.72	.5	.56	.68
Translation	.73	.35	.72	.66	.56	.7	.7	.51	.87	.61	.68	.78
Dictation	.61	.32	.76	.7	.49	.71	.68	.73	.73	.71	.65	.77
AC	.65	.47	.63	.58	.52	.65	.67	.55	.71	.54	.64	.69
Total	.74	.43	.79	.7	.56	.75	.75	.64	.86	.66	.71	.82
First Attempt Avs.												
Grammar	.63	.39	.7	.52	.5	.64	.59	.58	.75	.49	.53	.67
Translation	.74	.42	.83	.64	.53	.75	.72	.63	.8	.59	.63	.75
Dictation	.6	.27	.77	.63	.52	.69	.64	.53	.71	.66	.58	.72
AC	.72	.54	.81	.69	.6	.79	.74	.66	.8	.72	.73	.82
Total	.72	.46	.83	.67	.58	.78	.73	.65	.82	.66	.67	.8
Time												
12 Week	-	.63	.9	.84	.73	.92	.95	.81	.77	.79	.92	.88
18 Week	.95	.61	.88	.88	.71	.91	-	.79	.86	.85	.91	.92

Figure III-G-5 - Correlation of Internal Scores and 12 and 18 Week MLA Scores for CAI

	Grade	Test	Class	AC	ACS	Dictation/ Grammar	Oral	Read	Write	Speak	Listen	Total
12 Week Variables												
Grade	-	.97	.86	.87	.9	.93	.74	.54	.89	.89	.77	.92
Test	.97	-	.96	.8	.85	.91	.83	.53	.87	.88	.76	.91
Class	.86	.96	-	.64	.72	.82	.87	.47	.78	.78	.71	.81
Subtests												
AC	.87	.8	.64	-	.73	.72	.56	.65	.81	.82	.68	.86
ACS	.9	.85	.72	.73	-	.79	.57	.48	.75	.81	.72	.81
Dictation/Grammar	.93	.91	.82	.72	.79	-	.65	.39	.85	.78	.68	.82
Oral	.74	.83	.87	.56	.57	.65	-	.33	.63	.71	.56	.67
MLA												
Read	.54	.53	.47	.65	.48	.39	.33	-	.55	.58	.57	.72
Write	.89	.87	.78	.81	.75	.85	.63	.55	-	.83	.69	.94
Speak	.89	.88	.78	.82	.81	.78	.71	.58	.83	-	.7	.93
Listen	.77	.76	.71	.68	.72	.68	.56	.57	.69	.7	-	.82
Total	.92	.91	.81	.86	.81	.82	.67	.72	.94	.93	.82	-

Figure III-G-6 - Correlation of 12 Week DLI Grades and MLA Form MA Scores for CAI

	Grade	Test	Class	AC	ACS	Dictation	Grammar	Oral	Read	Write	Speak	Listen	Total
18 Week Variables													
Grade	-	.98	.97	.77	.78	.89	.83	.7	.72	.87	.85	.82	.93
Test	.98	-	.91	.79	.8	.91	.83	.72	.69	.82	.83	.78	.89
Class	.97	.91	-	.69	.72	.81	.8	.62	.71	.89	.83	.82	.93
Subtests													
AC	.77	.79	.69	-	.55	.7	.5	.51	.51	.62	.56	.59	.64
ACS	.78	.8	.72	.55	-	.7	.52	.61	.62	.71	.79	.79	.82
Dictation	.89	.91	.81	.7	.7	-	.62	.52	.46	.78	.79	.71	.81
Grammar	.83	.83	.8	.5	.52	.62	-	.62	.65	.68	.62	.52	.7
Oral	.7	.72	.62	.51	.61	.52	.62	-	.76	.41	.64	.68	.65
MLA													
Read	.72	.69	.71	.51	.62	.46	.65	.76	-	.58	.68	.69	.77
Write	.87	.82	.89	.62	.71	.78	.68	.41	.58	-	.76	.78	.91
Speak	.85	.83	.83	.56	.79	.79	.62	.64	.68	.76	-	.84	.93
Listen	.82	.78	.82	.59	.79	.71	.52	.68	.69	.78	.84	-	.91
Total	.93	.89	.93	.64	.82	.81	.7	.65	.77	.91	.93	.91	-

Figure III-G-7 - Correlation of 18 Week DLI Grades and MLA Form MB Scores for CAI

	Overall Grade (12)	Test (12)	Class Grade (12)	ACS (12)	AC (12)	Dictation Grammar (12)	Oral (12)	Overall Grade (18)	Test (18)	Class Grade (18)	ACS (18)	AC (18)	Dictation (18)	Grammar (18)	Oral (18)
12 Week Grades															
Overall Grade	-	.98	.97	.81	.84	.79	.85	.94	.91	.9	.77	.73	.77	.56	.91
Test	.98	-	.9	.87	.84	.78	.86	.89	.89	.84	.75	.74	.75	.54	.87
Class Grade	.97	.9	-	.7	.79	.77	.8	.94	.89	.93	.75	.67	.75	.56	.9
ACS	.81	.87	.7	-	.66	.54	.62	.7	.75	.61	.76	.71	.67	.32	.67
AC	.84	.84	.79	.66	-	.56	.69	.75	.77	.7	.64	.73	.62	.4	.81
Dictation Grammar	.79	.78	.77	.54	.56	-	.58	.83	.75	.85	.48	.57	.61	.6	.78
Oral	.85	.86	.8	.62	.69	.58	-	.74	.73	.72	.61	.48	.64	.51	.75
18 Week Grades															
Overall Grade	.94	.89	.94	.7	.75	.83	.74	-	.96	.97	.72	.73	.84	.7	.91
Test	.91	.89	.89	.75	.77	.75	.73	.96	-	.87	.76	.79	.89	.7	.89
Class Grade	.9	.84	.93	.61	.7	.85	.72	.97	.87	-	.63	.65	.74	.64	.88
ACS	.77	.75	.75	.76	.64	.48	.61	.72	.76	.63	-	.59	.65	.3	.69
AC	.73	.74	.67	.71	.73	.57	.48	.73	.79	.65	.59	-	.68	.37	.65
Dictation	.77	.75	.75	.67	.62	.61	.64	.84	.89	.74	.65	.68	-	.47	.79
Grammar	.56	.54	.56	.32	.4	.6	.51	.7	.7	.64	.3	.37	.47	-	.52
Oral	.91	.87	.9	.67	.81	.78	.75	.91	.89	.88	.69	.65	.79	.52	-

Figure III-G-8 - Intercorrelation of 12 and 18 Week DLI Grades for non-CAI

	Read (12)	Write (12)	Speak (12)	Listen (12)	Total (12)	Read (18)	Write (18)	Speak (18)	Listen (18)	Total (18)
12 Week MLA										
Read	-	0.32	0.31	0.45	0.6	0.57	0.24	0.39	0.36	0.45
Write	0.32	-	0.58	0.37	0.84	0.39	0.73	0.44	0.49	0.7
Speak	0.31	0.58	-	0.53	0.83	0.49	0.43	0.59	0.68	0.66
Listen	0.45	0.37	0.53	-	0.69	0.54	0.32	0.59	0.47	0.57
Total	0.6	0.84	0.83	0.69	-	0.62	0.64	0.65	0.67	0.8
18 Week MLA										
Read	0.57	0.39	0.49	0.54	0.62	-	0.31	0.5	0.48	0.64
Write	0.24	0.73	0.43	0.32	0.64	0.31	-	0.53	0.55	0.85
Speak	0.39	0.44	0.59	0.59	0.65	0.5	0.53	-	0.66	0.83
Listen	0.36	0.49	0.68	0.47	0.67	0.48	0.55	0.66	-	0.81
Total	0.45	0.7	0.66	0.57	0.8	0.64	0.85	0.83	0.81	-

Figure III-G-9 - Intercorrelation of Form MA and Form MB MLA Scores for non-CAI

	Read (12)	Write (12)	Speak (12)	Listen (12)	Total (12)	Read (18)	Write (18)	Speak (18)	Listen (18)	Total (18)
12 Week Grades										
Overall Grade	.32	.74	.73	.57	.82	.44	.66	.65	.81	.81
Test	.3	.69	.71	.6	.79	.41	.61	.61	.81	.77
Class Grade	.32	.77	.72	.51	.81	.46	.68	.67	.77	.82
Subtests										
ACS	.3	.49	.65	.61	.67	.3	.46	.53	.72	.72
AC	.28	.57	.66	.56	.7	.35	.5	.63	.84	.71
Dictation/Grammar	.22	.82	.58	.37	.73	.48	.74	.51	.62	.77
Oral	.2	.51	.54	.47	.59	.26	.43	.45	.61	.54
18 Week Grades										
Overall Grade	.25	.79	.71	.52	.81	.44	.77	.69	.74	.86
Test	.26	.74	.68	.59	.79	.41	.75	.68	.75	.84
Class Grade	.22	.78	.69	.42	.77	.44	.72	.66	.69	.81
Subtests										
ACS	.23	.55	.51	.41	.6	.29	.53	.58	.66	.65
AC	.28	.54	.64	.58	.68	.4	.52	.6	.72	.69
Dictation	.2	.62	.61	.53	.68	.3	.71	.64	.58	.74
Grammar	.12	.56	.36	.37	.51	.29	.55	.3	.38	.52
Oral	.24	.71	.64	.51	.74	.39	.7	.73	.77	.83

Figure III-G-10 - Correlation of 12 and 18 Week DLI Grades with Form MA and MB MLA Scores for non-CAI

	ALAT	3Week
ALAT	-	.47
3 Week	.47	-
12 Week		
Grade	.53	.81
Test	.53	.78
Class	.48	.77
ACS	.42	.67
AC	.55	.62
Dictation/ Grammar	.4	.77
Oral	.51	.59
MLA/MA		
Read	.41	.57
Write	.63	.7
Speak	.53	.63
Listen	.3	.67
Total	.58	.74

	ALAT	3Week
18 Week		
Grade	.45	.75
Test	.47	.76
Class	.47	.76
ACS	.62	.61
AC	.19	.59
Grammar	.45	.66
Dictation	.32	.64
Oral	.5	.65
MLA/MB		
Read	.53	.6
Write	.43	.57
Speak	.58	.67
Listen	.47	.59
Total	.56	.68

	ALAT	3Week
12 Week Internal		
Grammar	.32	.61
Translation	.42	.48
Dictation	.38	.7
AC	.4	.59
Total	.41	.65
18 Week Internal		
Grammar	.23	.43
Translation	.35	.56
Dictation	.39	.59
AC	.33	.7
Total	.36	.63
First Attempt		
Grammar	.2	.49
Translation	.4	.51
Dictation	.36	.58
AC	.39	.64
Total	.37	.6

Correlation of pooled ALAT vs. 3 week test is .42

Figure III-G-11 - Correlation of ALAT and 3-Week Test with DLI Grades, MLA and Internal Scores for CAI

	ALAT	3Week		ALAT	3 Week
ALAT	-	.38			
3 Week	.38	-			
12 Week			18 Week		
Grade	.29	.76	Grade	.35	.73
Test	.19	.67	Test	.28	.65
Class	.38	.83	Class	.38	.77
ACS	.08	.46	ACS	.17	.58
AC	.17	.59	AC	.22	.44
Dictation/ Grammar	.4	.61	Grammar	.2	.41
			Dictation	.23	.56
Oral	.07	.66	Oral	.26	.73
MLA/MA			MLA/MB		
Read	.51	.12	Read	.49	.32
Write	.48	.59	Write	.47	.57
Speak	.3	.57	Speak	.18	.47
Listen	.1	.38	Listen	.26	.5
Total	.49	.6	Total	.46	.61

Figure III-G-12 - Correlation of ALAT and 3-Week Test with DLI Grades and MLA Scores for non-CAI

	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12	Week 13	Week 14	Week 15	Week 16	Week 17	Week 18
Week 7	-	.89	.89	.81	.79	.88	.83	.82	.44	.71	.88	.83
Week 8	.89	-	.88	.89	.88	.88	.86	.8	.5	.72	.87	.70
Week 9	.89	.88	-	.8	.77	.94	.95	.85	.57	.84	.93	.89
Week 10	.81	.89	.8	-	.91	.88	.83	.84	.49	.75	.83	.76
Week 11	.79	.88	.77	.91	-	.83	.79	.81	.53	.71	.77	.68
Week 12	.88	.88	.94	.88	.83	-	.95	.89	.58	.86	.9	.87
Week 13	.83	.86	.95	.83	.79	.95	-	.87	.61	.85	.89	.86
Week 14	.82	.8	.85	.84	.81	.89	.87	-	.44	.83	.91	.91
Week 15	.44	.5	.57	.49	.53	.58	.61	.44	-	.5	.48	.46
Week 16	.71	.72	.84	.75	.71	.86	.85	.83	.5	-	.83	.8
Week 17	.88	.87	.93	.83	.77	.9	.89	.91	.48	.83	-	.95
Week 18	.83	.79	.89	.76	.68	.87	.86	.91	.46	.8	.95	-

Figure III-G-13 - Intercorrelation of Weekly Time Averages

	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12	Week 13	Week 14	Week 15	Week 16	Week 17	Week 18
Week 7	-	.99	.97	.95	.92	.92	.91	.9	.85	.83	.87	.87
Week 8	.99	-	.99	.98	.96	.95	.94	.92	.88	.86	.89	.89
Week 9	.97	.99	-	.98	.96	.96	.96	.95	.91	.9	.93	.92
Week 10	.95	.98	.98	-	.99	.99	.98	.97	.93	.92	.94	.93
Week 11	.92	.96	.96	.99	-	.99	.98	.96	.93	.92	.93	.9
Week 12	.92	.95	.96	.99	.99	-	.99	.98	.96	.95	.95	.93
Week 13	.91	.94	.96	.98	.98	.99	-	.99	.97	.96	.97	.95
Week 14	.9	.92	.95	.97	.96	.98	.99	-	.96	.96	.98	.97
Week 15	.85	.88	.91	.93	.93	.96	.97	.96	-	.98	.97	.94
Week 16	.83	.86	.9	.92	.92	.95	.96	.96	.98	-	.99	.96
Week 17	.87	.89	.93	.94	.93	.95	.97	.98	.97	.99	-	.99
Week 18	.87	.89	.92	.93	.9	.93	.95	.97	.94	.96	.99	-

Figure III-G-14 - Intercorrelation of Weekly Time Recency Scores
(recency weight = 2)

Correlation of 12th Week Average Time and Test, Grade and MLA at 12 Weeks

	Test	Grade	MLA
Av. Time	.92	.93	.93

Correlation of 12th Week Time Recency and Test, Grade and MLA at 12 Weeks

	Test	Grade	MLA
Time Recency	.89	.92	.92

Correlation of 18th Week Average Time and Test, Grade and MLA at 18 Weeks

	Test	Grade	MLA
Av. Time	.8	.8	.84

Correlation of 18th Week Time Recency and Test, Grade and MLA at 18 Weeks

	Test	Grade	MLA
Time Recency	.86	.87	.92

Correlation of Overall Average Time with Test, Grade and MLA at 12 Weeks and Test, Grade and MLA at 18 Weeks

	12 Test	12 Grade	12 MLA	18 Test	18 Grade	18 MLA
Av. Time	.92	.92	.92	.88	.9	.92

Figure III-G-15

Section H

Part One of this appendix contains the questionnaire administered to the 27 surviving DLI CAI students upon completion of the operational phase. The results of the questionnaire are tabulated above the selection categories in each item.

Part Two contains the verbatum responses of the students to question 13, the general information question.

PART ONE

Questionnaire for DLI Students in CAI Section

This questionnaire is for the purpose of statistical analyses (without identifying individual students), and for gathering suggestions for improving the course. Your instructor will not see your responses.

1. Did you find typing an acceptable way of inputting foreign language messages? (circle one)

0	1	2	16	8
very poor	poor	o. k.	good	excellent

2. Did you find typing an acceptable way of receiving messages from the computer terminal?

0	1	3½	16½	6
very poor	poor	o. k.	good	excellent

3. Was the quality of tape messages good enough?

0	2	9	10	6
very poor	poor	o. k.	good	excellent

4. How reliable was the terminal equipment you used?

0	5	5	15	2
very poor	poor	o. k.	good	excellent

5. Did you find the exercises in the Language Laboratory confusing or misleading?

0	24	2	1
never	occasionally	frequently	very frequently

6. How helpful did you find the following exercise types: (check one for each exercise)

	poor	fair	all right	good	very good
Textual Substitution Transformation	0	1	9	16	1
Aurally Cued Substitution Transformation	0	3	7	16	1
Translation	6	4	4	11	2
Sentence Dictation	0	3	4	13	7
Name Dictation	7	5	9	5	1
Number Dictation	0	3	5	11	8
Aural Comprehension of a Discourse	2	3	7	5	10
Aural Comprehension of Sentences	3	2	6	6	10
Aurally Cued Transformation	0	3	10	12	2

7. How interesting did you find the exercise types to be? (check one for each exercise)

	poor	fair	all right	good	very good
Textual Substitution Transformation	2	5	13	7	0
Aurally Cued Substitution Transformation	3	7	11	6	0
Translation	5	8	5	7	2
Sentence Dictation	0	6	9	7	5
Name Dictation	7	7	6	3	4
Number Dictation	2	5	8	6	6
Aural Comprehension of Discourse	4	7	6	3	7
Aural Comprehension of Sentences	5	5	6	3	8
Aurally Cued Transformation	1	7	11	6	2

8. How would you compare the effectiveness of CAI as opposed to exercises done as homework?

0	2	3	10	12
much worse	worse	the same	better	much better

9. At what parts of the course would CAI be most useful?

0	11	2	2	12
never	beginning	middle	end	throughout

10. What is your overall impression of the potential usefulness of CAI as a language learning aid?

½	1	9½	11	5
no potential	not much	so-so	great	very great

11. How do you think you did as compared to non-CAI taught students?

0	8½	12½	5	0
much worse	worse	as well	better	much better

(Note: one student said it was too soon for him to know, and he therefore did not circle any answer.)

12. If you were now beginning Russian and had a choice between a section using CAI and a regular section, which would you choose? (circle one)

7	19	1	0
regular section	CAI section	don't care	don't know

13. Do you have any suggestions as to how the CAI course might be improved through additions, deletions, or other modifications (use reverse side).

PART TWO

Student Responses from Questionnaire for DLI Students in CAI

1. At first I was rather dubious as to the usefulness of the computer assisted method, however, after getting adjusted to the machine I concluded the following additions, deletions, etc. which may enhance this method of teaching a language:

   1) More aural comprehension of paragraphs and sentences.

   2) As compared with the regular, non CAI students I found my comprehension to be below my compatriots in section A.

   3) As the course progressed the lessons naturally became more difficult and I think it wise if the lessons would be geared to the words, grammar, etc. learned the previous day and NOT for the next day. I say this because it caused me no ending anguish to have the computer "-----" my inadequacy of the language; as is, I had to study just as much that evening to get the words and grammar through my cerebrum. By messing up one, say aural transformation of sentences, I would mess up the next (say sentence dictation) through anguish at myself.

All in all, I did find the computer a helpful device, not intending to be hypocritical, in learning Russian.

It also was useful in breaking up the day which I may have found "Dovol'No Skuchno". I have had a computer class in college and see it is the coming thing in teaching the masses.

---

2. Additional dictations seem necessary for CAI class--the transition to regular class may prove to be traumatic.

---

3. The translations well could have been deleted. The computer might be tried for testing purposes as well as for the uses to which it was employed during this experiment.

As far as mechanics involved, it is occasionally frustrating being at the mercy of cantankerous long distance phone lines.

For the last few weeks, having to attend the late session each evening dulled my individual performance. A student needs to be alert to effectively take advantage of computer teaching. This just isn't possible after 4 hours of class, 2 hours of study hall, and the time the Army life requires for inspections, morning details, and duties like server.

I hope the final results indicate definite advantages for a computer teaching program. I think a good use for the computer here at DLI in the future would be in a supplementary role to regular classes--for slower students, accentuated emphasis on more difficult parts of grammar, etc.--especially in place of supervised study hall.

---

4. CAI's greatest potential as far as I'm concerned, lies in teaching vocabulary, grammar, and structure of a language. Here I think it is very helpful. But, its weakness lies in the area of developing aural comprehension. This was the weakest part of the CAI program and definitely needs overhauling. That is why I feel CAI is not the best method for a course aimed primarily at aural comprehension. Use it to develop working knowledge of a language at a lower achievement level, then continue with an aural comprehension course.

CAI potential is good, in fact great, in areas of grammar, grammatical structure and vocabulary.

---

5. We need much more practice on dictation and comprehension. I feel this should be started much earlier and placed under more emphasis. Also it would be better if we were tested on the machines instead of taking the same tests as the rest of the class. They had a great advantage in writing practice which is a large part of the grade. We are not prepared for the dictation or aural comprehension parts of the 6 week tests. CAI does, I feel, have a lot of potential as far as learning language goes, grammar wise, but the emphasis in the regular course is based on comprehension.

---

6. The course, as of the present, is set up for regular type sections. Some adjustments need to be made in testing and evaluating CAI students which would allow for differences in the two programs.

I think the ideal situation would be the use of both types of lab. by all students. The regular lab. for aural comprehension and speaking practice and the CAI lab. for grammar and construction. Perhaps a schedule such as:

1st hour--dialogue presentation
4th hour--new grammar
5th hour--use of new grammar in CAI lab.
3rd hour--aural comprehension in reg. lab.
2nd hour--free conversation using material of lesson of the day
6th hour--new dialogue

---

7. CAI might work more effectively if speed could somehow be involved in the exercises. The major hangup is the fact that CAI are not graded according to the teaching methods employed for them, creating a basic inequality. As an experimental project I'm not sure it can prove or disprove anything. I personally think it was more detrimental for slow learners. Besides it is hard to prove an advantage for better students as quite probably they will show little effect one way or the other. The most advantageous use of CAI would be as an extra study device for poor grammar students for it definitely helps in this area. Typing is one definite

advantage I derived from the program; but the testing methods employed for all students is so vastly different from the computer. Aural appears to be the big weakness of the program, grammar is the strong point. Unfortunately in this program one doesn't hear enough Russian spoken, which contributes to the feeling of inadequacy when hearing spoken Russian during a major test. Quite simply I have no EOB "the crippler" button to push. Seeing mistakes are fine but sometimes it takes more than seeing mistakes. You have to hear Russian to speak it.

---

8. I think that if CAI was used that the students should be graded different and take different tests than the students who aren't taking. That is the basis for my last answer. Go all the way with one thing or not at all. Switching back and forth is no good.

Also I think the translation blocks should be eliminated. The main purpose of the course is comprehension. They took up too much time and I don't really think they helped that much. Maybe it would have been simpler and sentences we were familiar with from our dialogue. This would have achieved better results as well as helping teach the dialogues.

I also felt a lot of times that it wasn't quite clear what the machine wanted. Maybe instructions could be put in instead of samples only. The models were sometimes misleading.

As far as grammar goes, the machine was great but for aural comprehension it is a poor substitute.

---

9. 1. More aural training, dictation, comprehension of sentences.
   2. Better coordination of review lessons and previous class lessons.
   3. Grading should be separate from rest of class.

---

10. Deletion of frequency of name recognition--more use of aural comprehension; i.e. sentences and discourses--greater flexibility of interchangeable words and phrases--higher degree of individual program determination according to individaul needs.

---

11. Shorter sentences in most all exercises to avoid mistakes because of length. Closer parallel to material already learned. Several times we've had exercises or phrases we have not learned. Also more comprehension and dictation with perhaps a more responsive oral system.

---

12. More aural comprehension--less aural and substitution transformation.

---

13. Less breakdowns (lines dropping).

---

14. Name recognition, unless improved, not of much use or benefit. Minor mistakes such as missing periods, etc. should not have to be retyped.

---

15. 1. Better tapes.
    2. Elimination of line drops.

---

16. 1. More sentence dictation
    2. More and longer aural comprehension of discourse
    3. Presence of teacher to explain points of grammar
    4. Have CAI exercises covering the lesson of the previous day, rather than the newly presented material.
    5. Allow students to do extra work on any particular exercise (if he feels it is necessary and has extra time available.)

---

17. 1. A differentiation between typing errors and spelling errors if possible, would be very helpful
    2. Elimination of the EOB to prevent the "easy way out". Should type all threee times in an honest attempt and then be given the answer.
    3. Dictations of all types should be given faster and repetitions should be held to a maximum of three times.
    4. More aural comprehension of discourse and sentences.

---

18. Translations have too many possibilites to be programmed adequately. How about using the exercises as review instead of learning (hah!) tomorrow's grammar and vocabulary in this manner. Better aural comprehension than name identification (becomes redundant after awhile).

---

19. Vocabulary should not be taken from what was presented that day. Item recognition and vocabulary dictation should be deleted. Answers accepted for translations should be more flexible. There should be more aural comp. of sentences and less aural comp. of discourse. Some of the aural cue transformations were too long to be effective; i.e. it was hard just remembering the sentence. When an item is added or changed it should be explained to all students before letting them start. Some means should be found to lower the amount of frustration from dropped lines, machines demanding wrong answer, use of vocabulary we haven't had, vagueness etc.

---

20. More aural comprehension of sentences. More dictation.

---

21. Translations should be discontinued (or cut way down after first few weeks) as, in general, there are way too many possible answers and they are not programmed. If a student is going to be allowed to skip exercises because he maintains a high average in that particular category then I think that this should be determined by each

individual exercise and not a general area (e.g., dictation includes sentences, numbers and names. A student may never make a mistake in numbers and names but make sufficient mistakes in sentences such that the has to do all the exercises again. Why waste time by repeating something that the student knows very well just because he happens to be a little weak in one respect). It seems to be virtually impossible to beat the machine (especailly on translations) and the student gets very discouraged and doesn't really want to try as he knows that it does him no good to try. Also give the student a way that he can skip listening to the tape again when not necessary as it is time consuming. Some exercises seem to be very strict in what the computer will accept, while at other times it will accept anything. Knowing how to type would help enormously as for myself, at least 90% of mistakes were from typing. If the keyboard were strictly in Russian it would not be as difficult because some letters are very easily mistyped (e.g., B for б, P for п etc,)

---

22. One of the most serious shortcomings of this method is that we are poorly prepared for competetion with the other sections in taking dictations. We should have a series of sentences dictated to us which would have to be typed in some time limit. Using the computer puts us under less strain with the present sentence dictation exercise because we have all the time we need and can always have the sentence played back. Thus we are poorly prepared for

6 week test dictations. There should be more emphasis on accuracy and less on speed in finishing the lesson. This could be accomplished by (machine) grading of all parts of the lesson.

---

23. My biggest complaint with the program was the 2-4 o'clock class. In order for the time I spent to be effective, I had to have had familiarized myself with the new grammar and vocabulary first, and there wasn't enough time when I had the 2 o'clock class to learn the material and then be able to do the exercises without looking in the book for every other answer.

If the purpose of the CAI program is to reenforce what is learned as well as teach, I feel the program could be improved by assigning the new grammar, etc. one day and doing the CAI exercises for that lesson the following day.

I thought the translation exercises were boring, unfruitful, and a waste of time, and I though the NAME DICTATIONS were not very helpful. The aural comprehension of sentences was not used often enough to be useful. In a course like this with emphasis on listening and understanding; I think the aural comprehension of sentence exercises should be included every day--perhaps in place of translations.

It's impossible for me to say now how well I did compared with the non-CAI part of the class YET. But one of the best advantages of the experiment is that it breaks up the monotony of 6 hours of class a day, makes the course interesting, and keeps up interest in the language.

---

24. A) Delete translation because:
    1. too many possibilities which are not programmed
    2. unprogrammed, but correct, answer is frustrating
    3. commas and question mark deletions when corrected by the machine are time consuming and drab

    B) Greater use of repetition in exercises such as textual subs., trans. aural cue subs., trans., aural trans.
    1. Induction requires more than one exampel of any particular grammatical point.
    2. Greater space and time could be afforded for this suggestion through exercise of suggestion "A".

    C) Give the student the option to repeat the tape message; in the aural drills a misplaced finger often is the cause for a repeated message which can be quite annoying. Again, a grammatical error is not corrected through the repetition of a word in an "ACS" exercise.

    D) Implement some method whereby the element of competition is removed from the experiment--i.e., booths, etc.

E) Stress "B" above--I feel the failure to recognize this nees is extremely detrimental to the future of the method.

---

25. No comment.

---

26. No comment.

---

27. No comment.

Section J

This section contains

1) a presentation of formuli used to compute internal coefficients of test reliability
2) tables of test reliabilities for DLI, MLA, and CAI scores.

Also presented are four graphs comparing the predictive power of various blocking variables as applied to major criterion measures. For example, Fig. III-J-3 was constructed by blocking the CAI section into quartiles by six different predictor measures (i.e., ALAT, 3-week test, 18-week MLA, CAI internal, CAI first attempt, time) and by then taking and plotting the various quartiles mean for the 12-week test.

The reliability r of a test may be thought of as its self-correlation in the test-retest sense. It may by extension then be considered as the correlation between scores on two "parallel" versions of the test. If one thinks of an observed score as the sum of a true score t and an error score E. then

$$r = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_E^2} = \frac{\sigma_t^2}{\sigma^2} \qquad \text{III-J-1}$$

where σ_t^2 is the true score variance. σ_E^2 the error variance, and σ^2 the total variance of the observed score.

If S is a composite score obtained as the sum of k scores $s_1, \ldots s_k$ on k parallel tests of reliability r, then the coefficient of correlation between S and one of its components s is:

$$(S,s) = \left(\frac{1 + (k-1)r}{k}\right)^{\frac{1}{2}} \qquad \text{III-J-2}$$

The reliability R of the composite S is:

$$R = \frac{kr}{1 + (k-1)r} \qquad \text{(Spearman-Brown formula) III-J-3}$$

The correlation coefficient (s.w) of two sets of independent scores s^i, w^i having individual reliabilities r_s, r_w is

$$(s,w) = \left(r_s r_w\right)^{1/2} \qquad \text{III-J-4}$$

where these reliabilities refer to the use of the two tests as independent measures of the *quality that they both measure*. Thus the *parallel test* reliability depends on what quality one is interested in measuring whereas the test-retest reliability is a property of the test itself. For example, the MLA test in listening may have a test-retest reliability of .90 *as a test of listening* and may at the same time have a reliability of only .50 considered *as a test of general language skill*.

The data we have are not adequate to determine the reliabilities of all the component tests individually. However, they do suffice to permit several estimates of most of the *average* or *composite* grades, considered then as *measures of general language skill as defined by the DLI academic grades*.

For a composite grade we can obtain two internal estimates of reliability; one will be the average intercorrelation of the components with one another; the second is obtained by obtaining the average intercorrelation of the components with the composite itself and then estimating an average r from this value from III-J-2.

The reliability of each composite DLI grade can then be estimated from that of its components from III-J-3. Finally the reliabilities of the various main tests can be obtained from their coefficients of correlations with the main DLI grades through (III-J-4).

	MLA (tot)	DLI (test)	DLI (class)	DLI (gde)	CAI (time)	CAI (int)	CAI (1st att)
DLI test		-	.88	not ind	.90	.81	.71
DLI class			-	not ind	.87	.81	.83
DLI gde					.90	.84	.78
MLA tot	.80	.86	.84	.88	.92	.82	.80
CAI time					-	.78	.73
CAI int						-	.91
CAI 1st att							-

Table III-J-1

Mean correlations of principal composite scores. Test reliabilities (general language skill).

	R of composite (internal measure)	R of composite (ref to DLI grade)
DLI (test)	.87	.89
DLI (class)	.87	.89
DLI (grade)	.94	.94
MLA (total)	.80	>.84
CAI (time)	.92	>.90
CAI (int)	.92	>.75
CAI (tst att)	.95	>.77

Table III-J-2

Estimated reliabilities of composite scores.

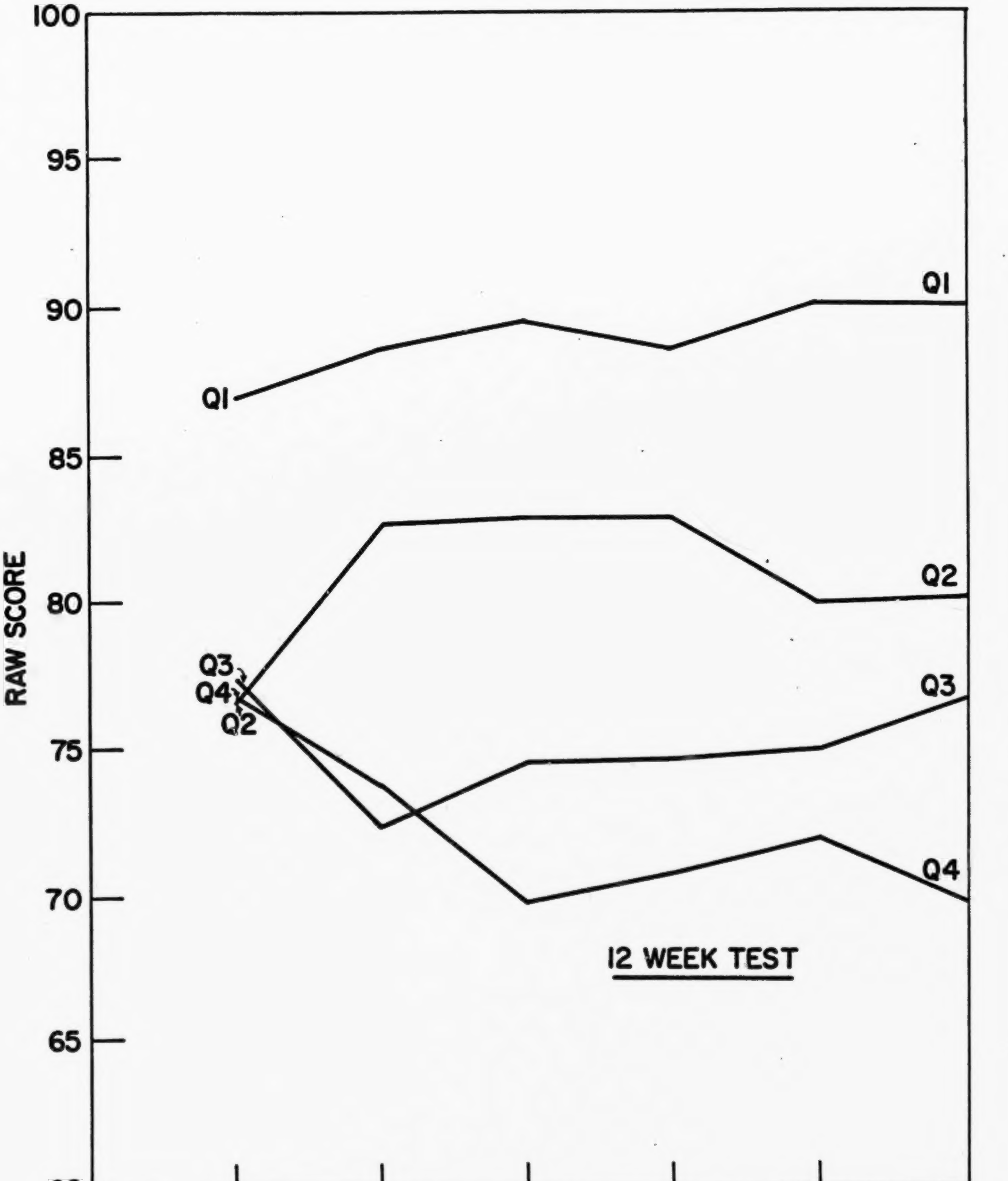


Figure III-J-3

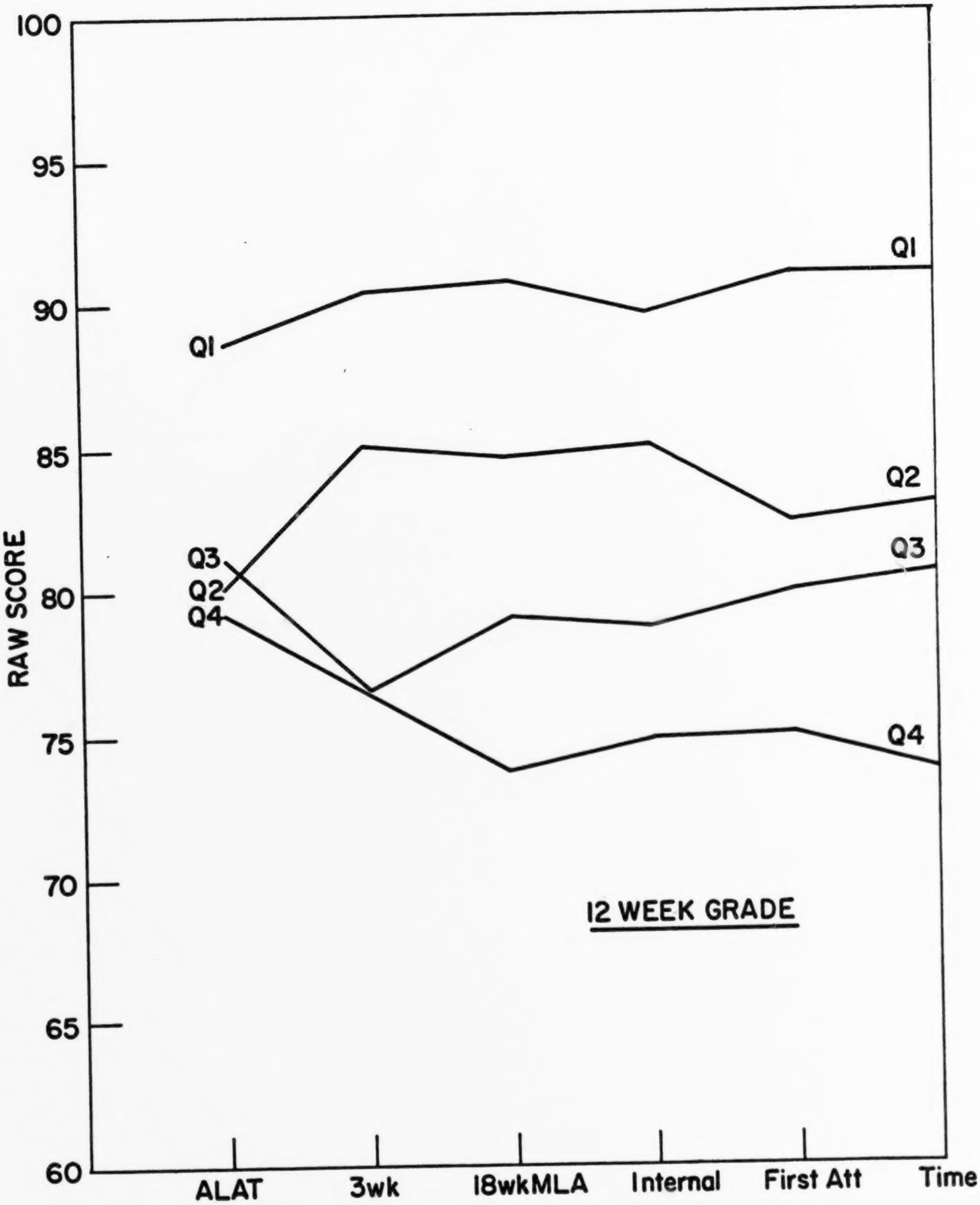


Figure III-J-4

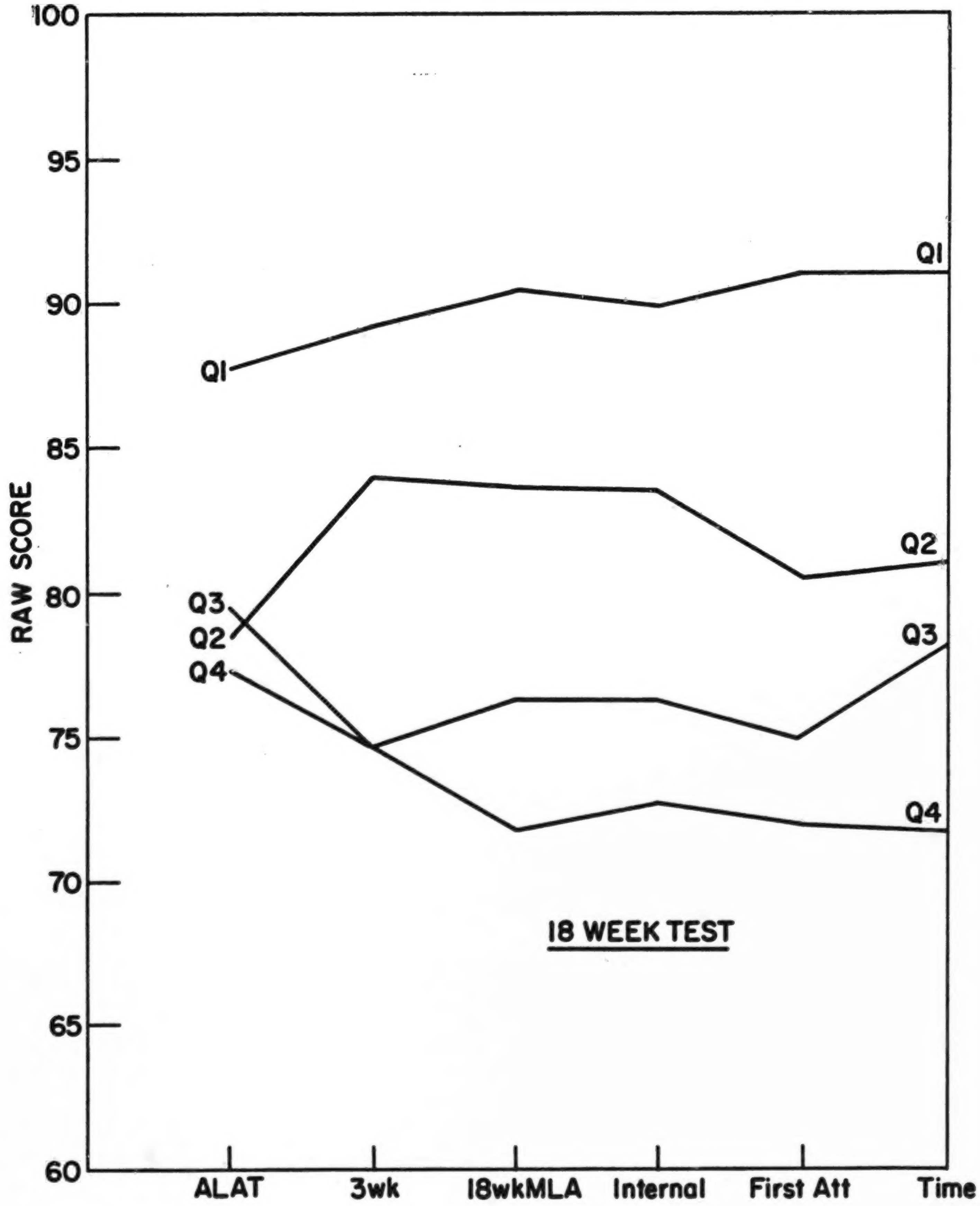


Figure III-J-5

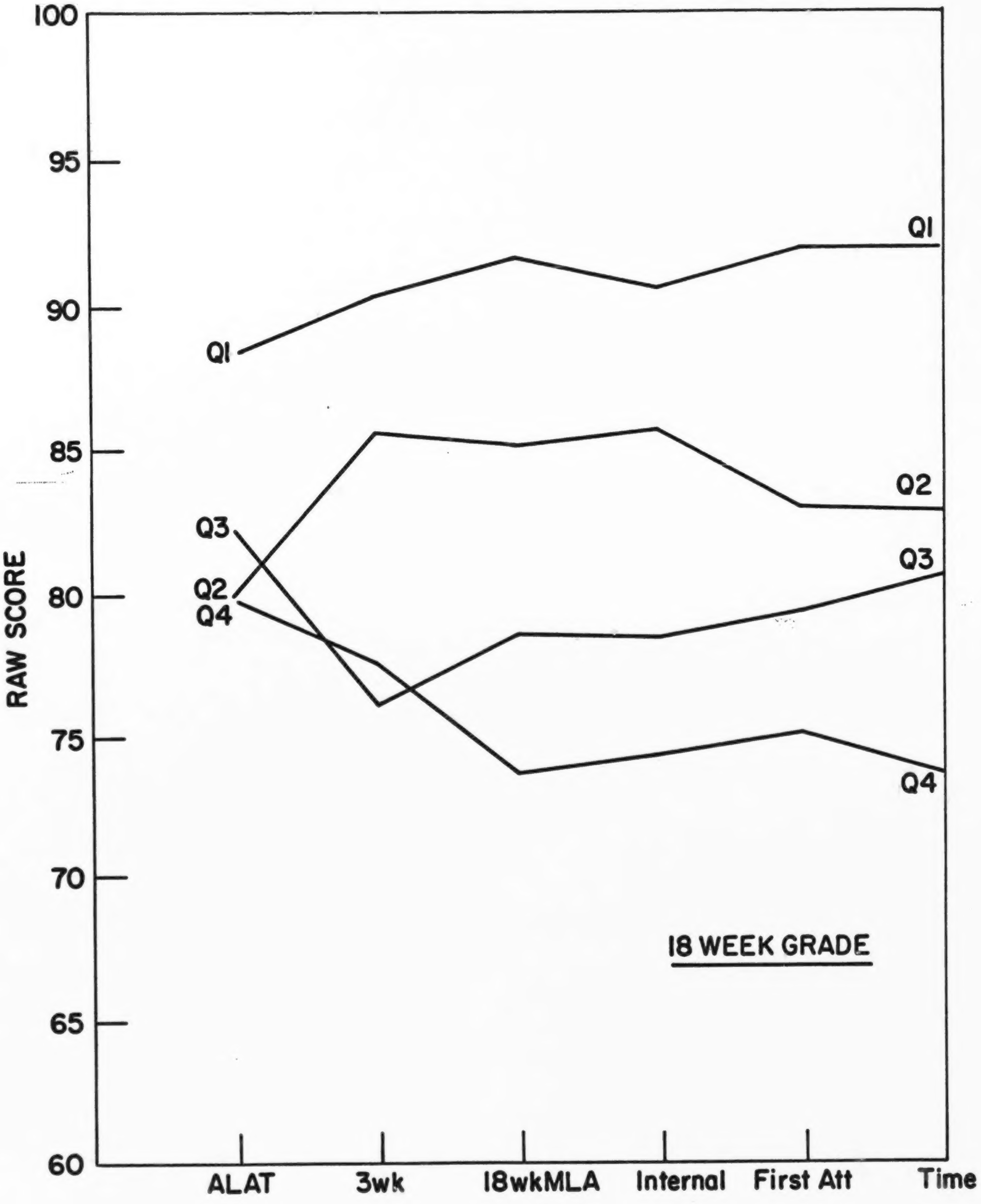


Figure III-J-6

Section K

This section contains performance comparisons for CAI and non-CAI sections on the tests and sub-tests administered after 24 weeks of instruction. In the interim period between the 18 and 24 week tests, two CAI students left the AC course for administrative, i.e., non-academic, reasons. Consequently, scores for these two students and their matched control counterparts do not appear in the graphs. The number of groups is taken as 28 rather than 30.

Also in this section is a rank order distribution (Fig. III-K-8) of the percentage of total available CAI time actually used by CAI students.

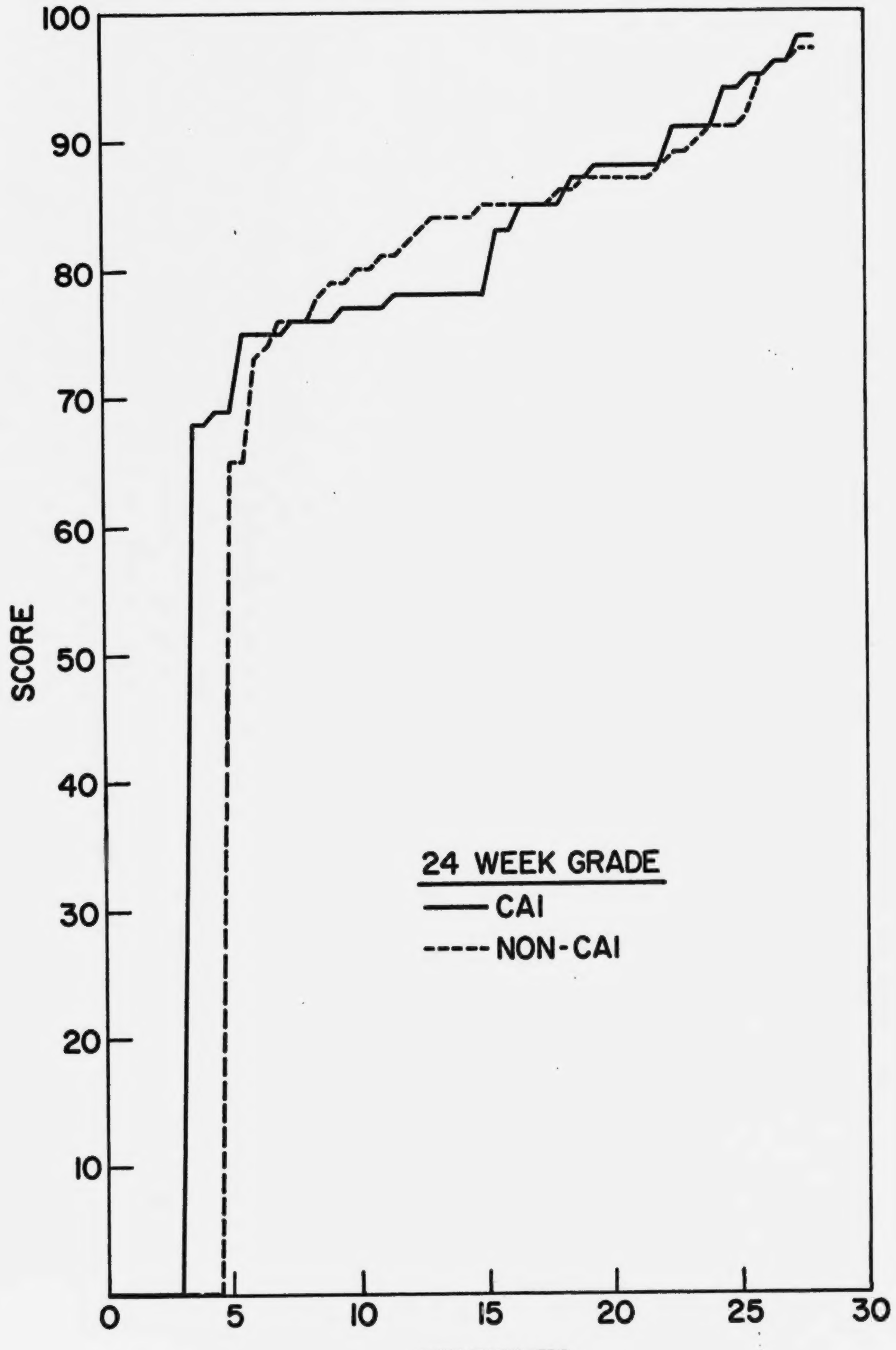


Figure III-K-1

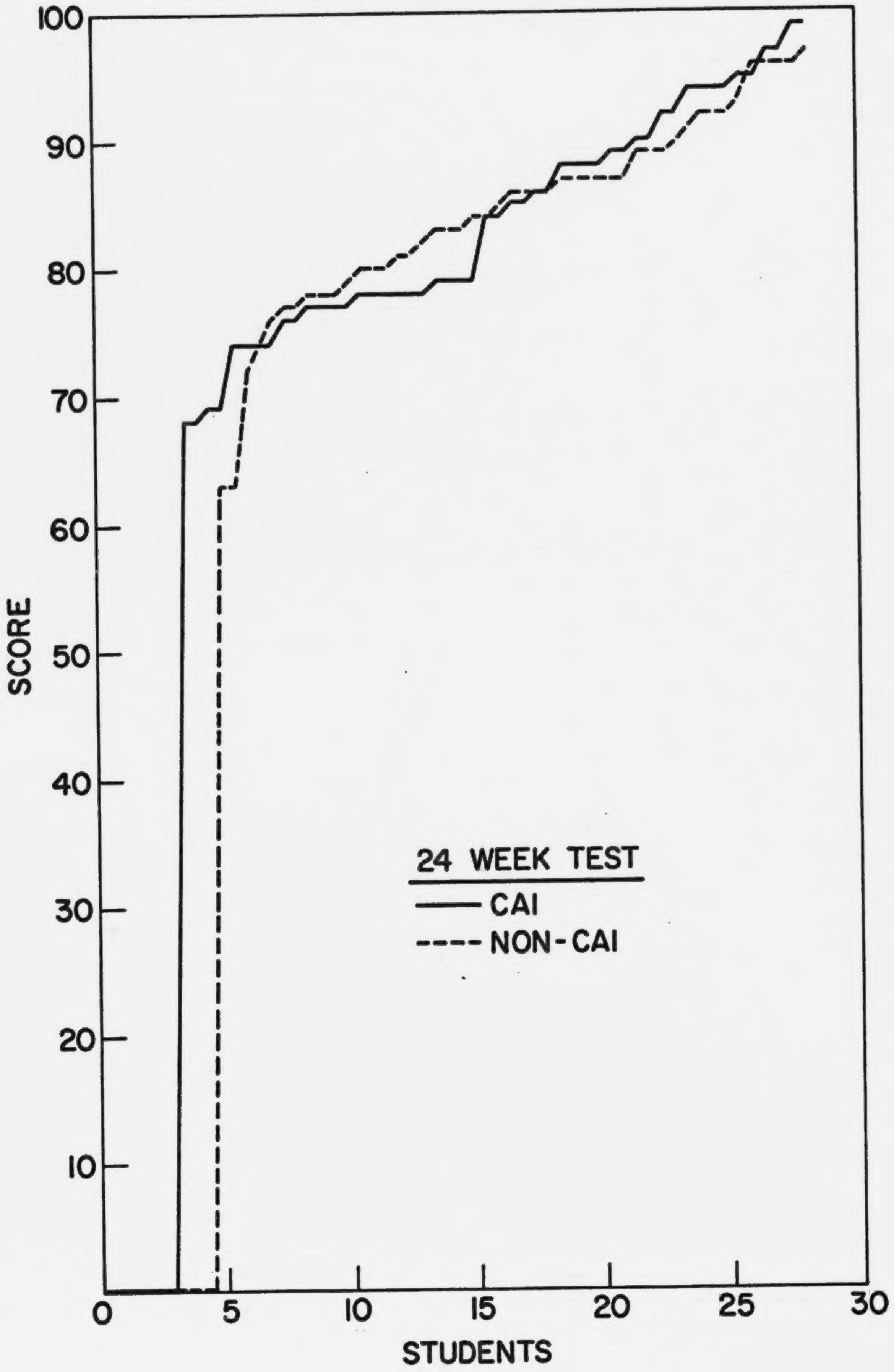


Figure III-K-2

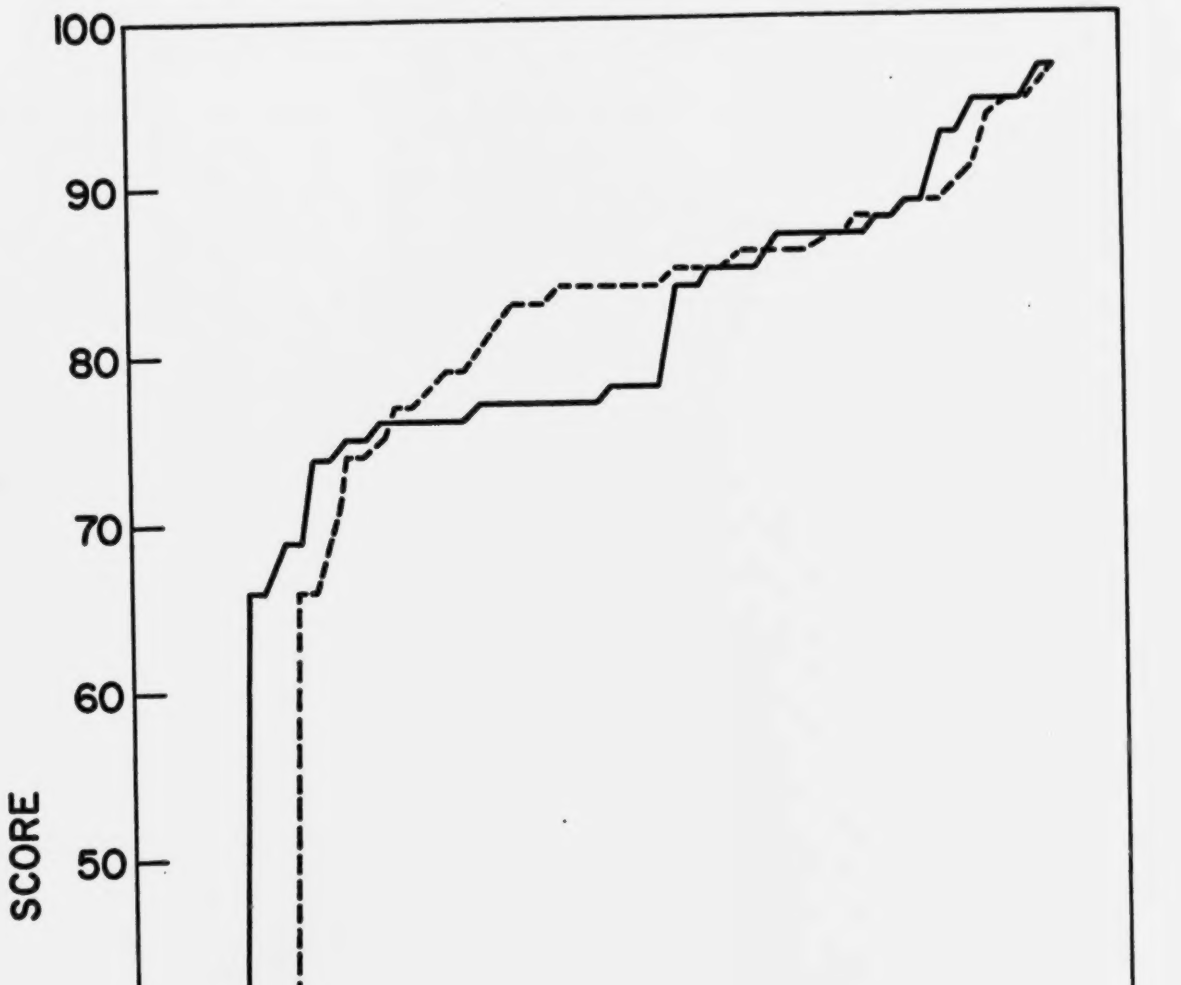


Figure III-K-3

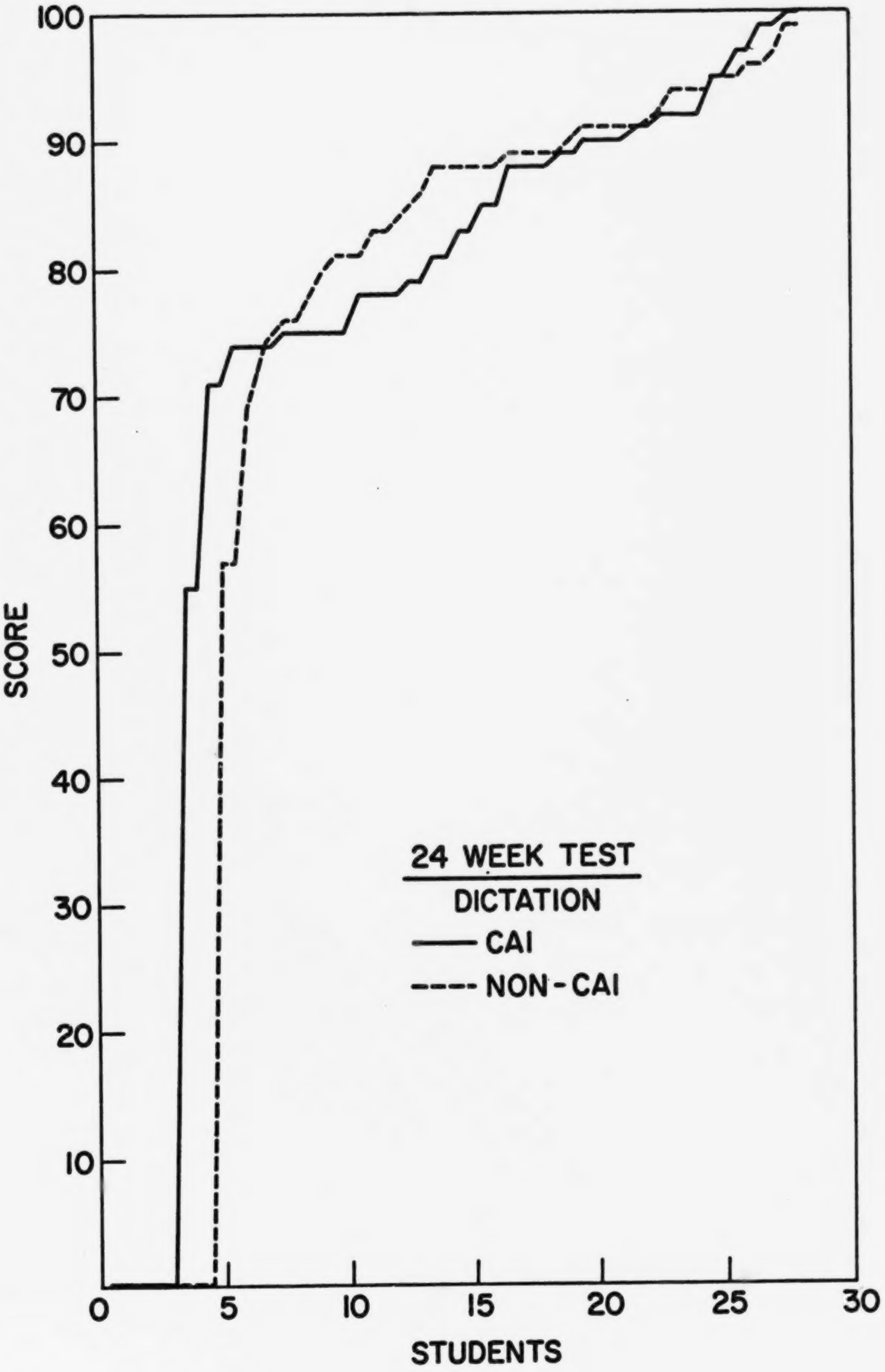


Figure III-K-4

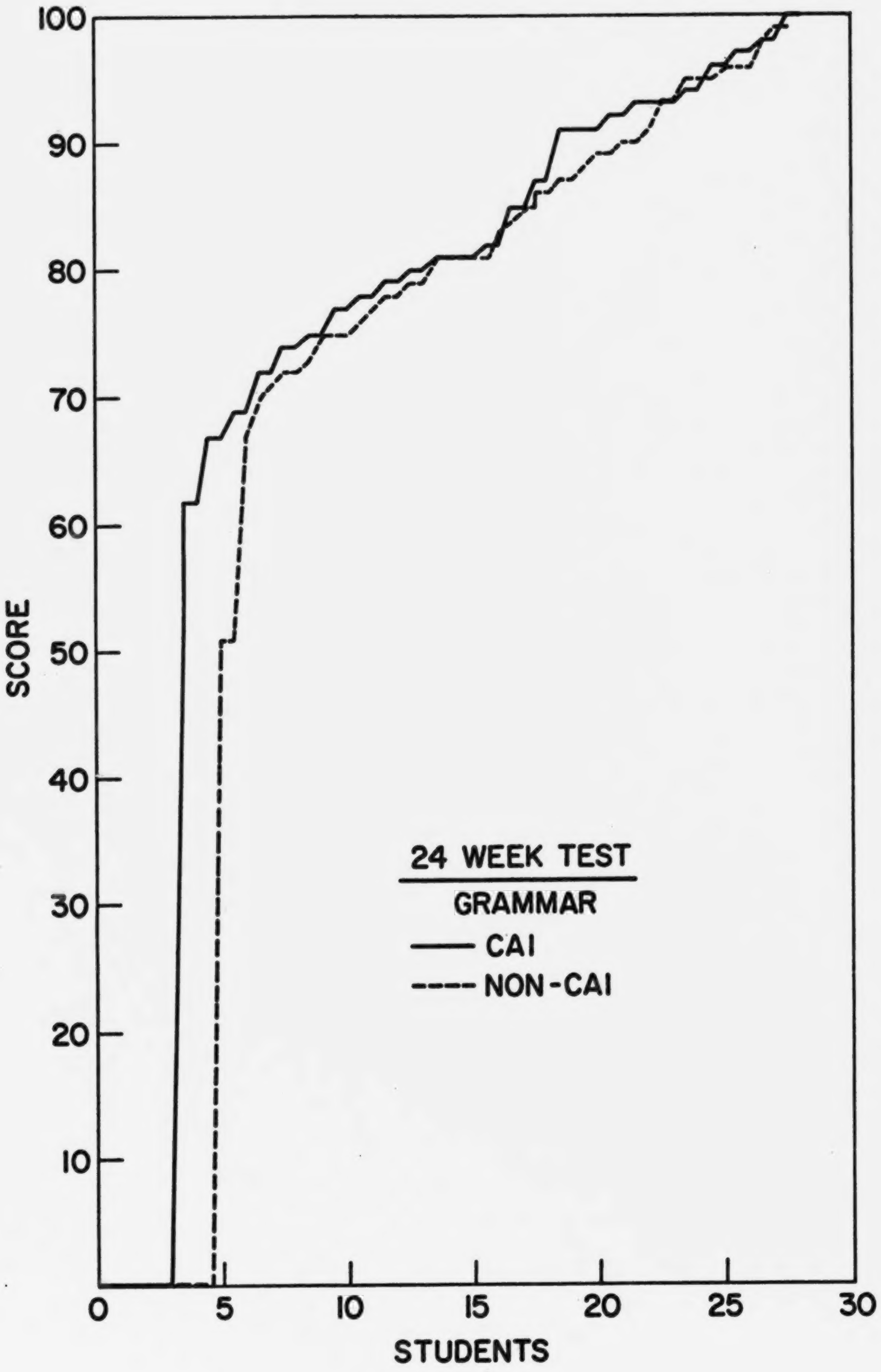


Figure III-K-5

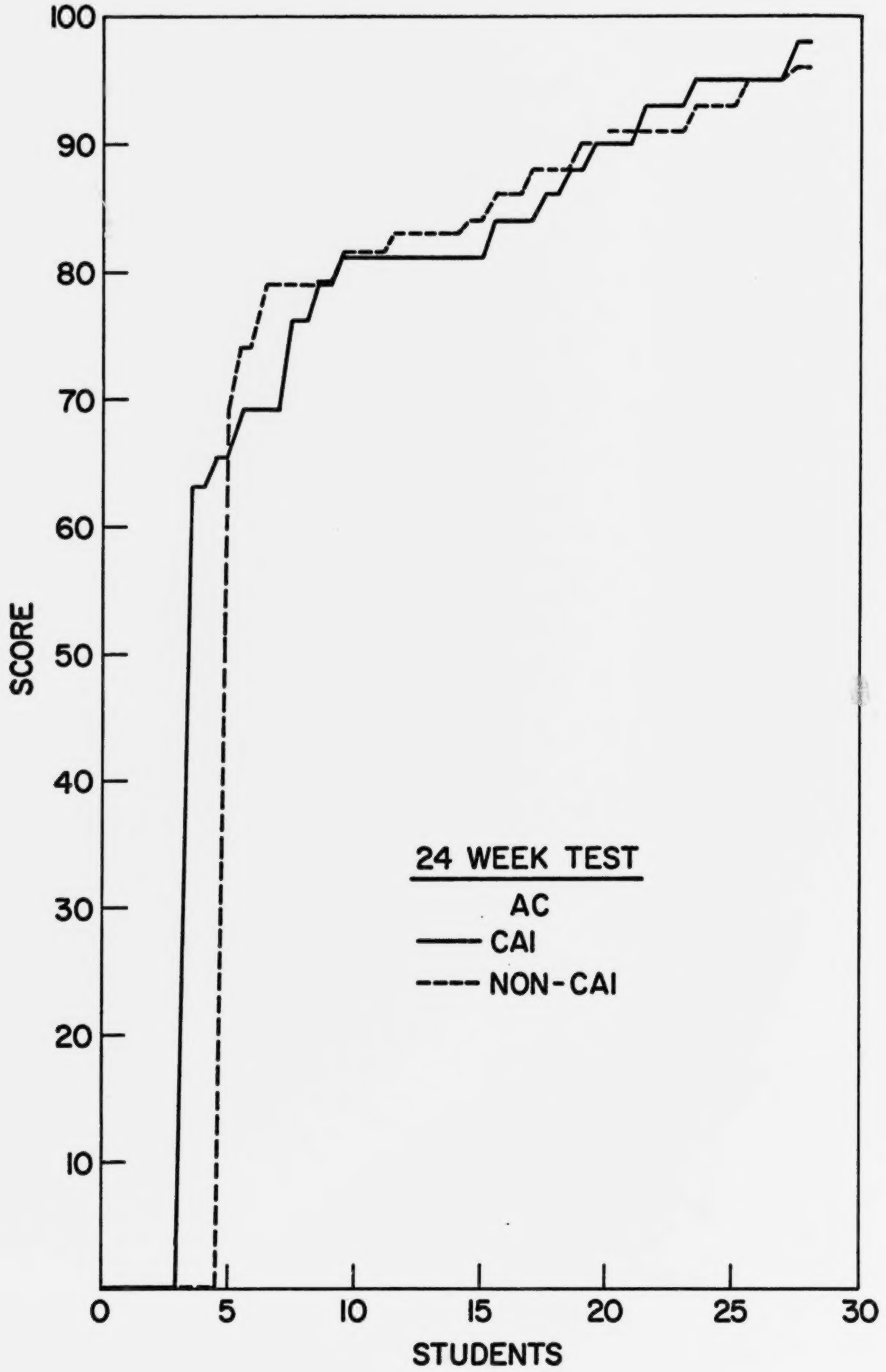


Figure III-K-6

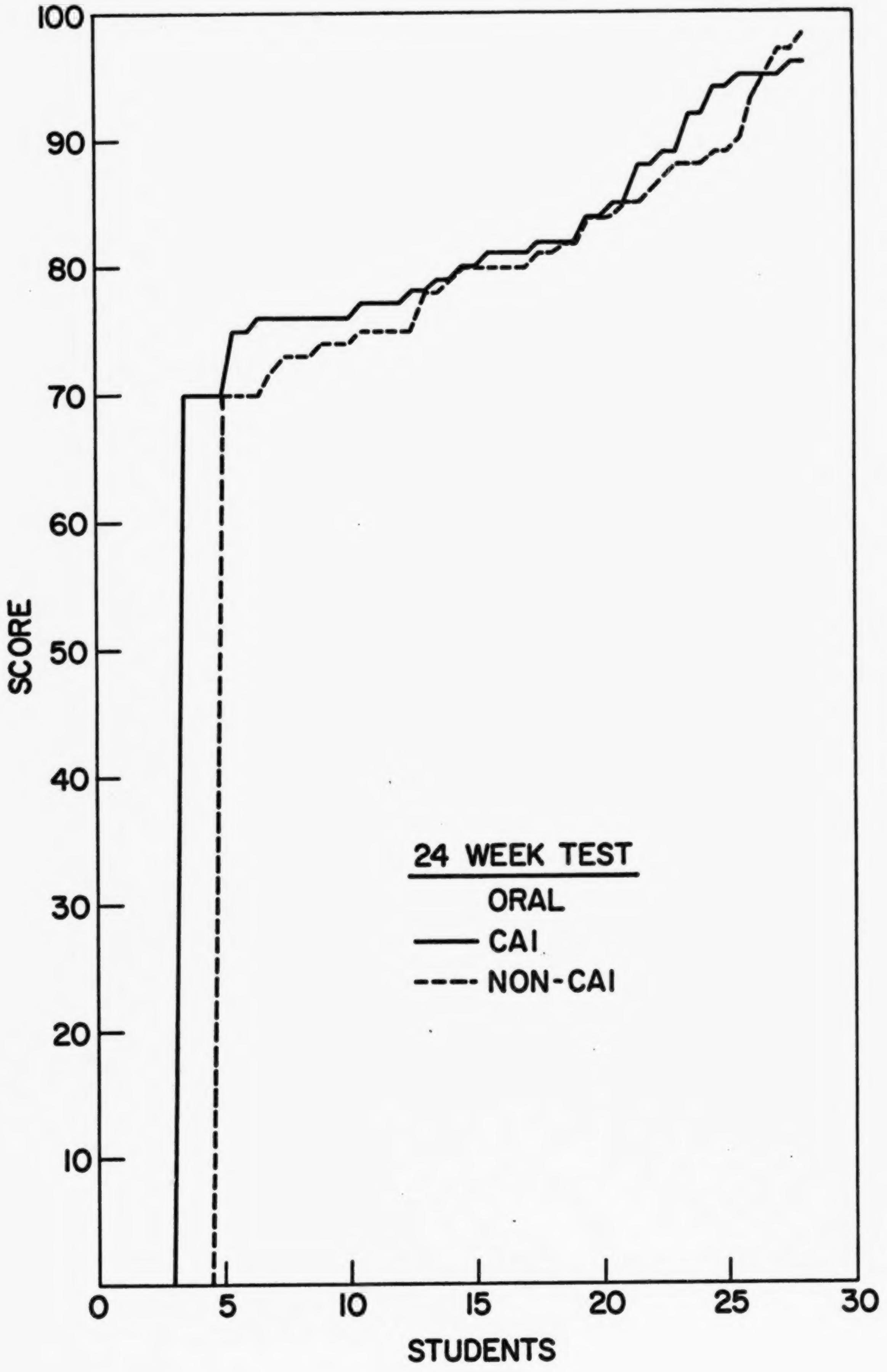


Figure III-K-7

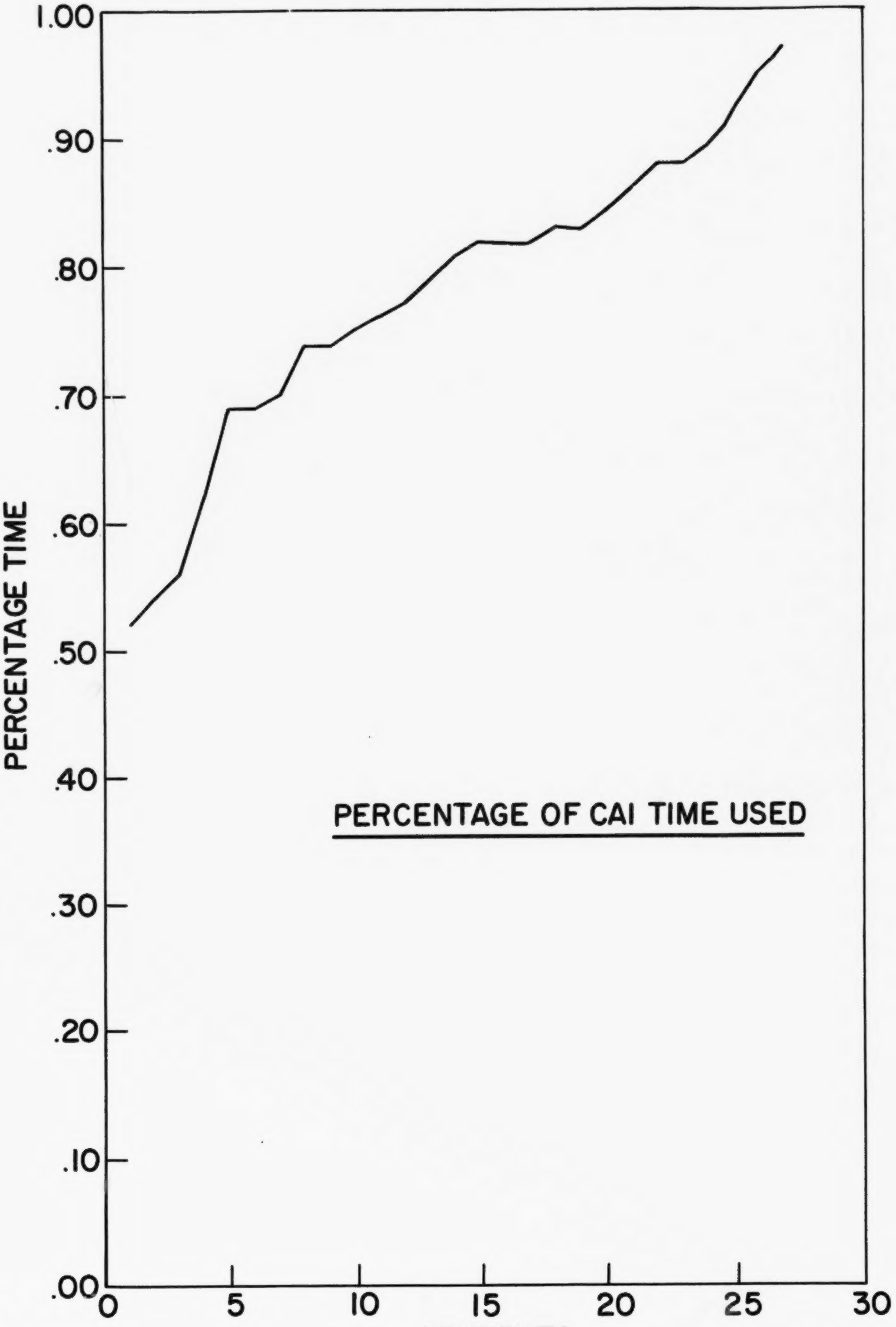


Figure III-K-8

Final Performance of Russian ACC Students

E. N. Adams

In the last week of August 1969 at the end of the Russian ACC course several measures of final student performance were obtained and compared to look for possible differences of retention between CAI and non-CAI students. We report here the comparison of performance on three final measures (1) the final academic average, (2) the Army Language Proficiency Test (ALPT) of listening comprehension and (3) of reading comprehension.

The student scores are presented in Table L-1 and plotted in graphs L-1, L-2, and L-3. In these graphs only 28 of the original 30 matched groups are shown, because CAI students No. 1 and 27 were removed from the course for administrative reasons other than academic performance.

The graphs show that the distributions of scores for CAI and non-CAI students are quite similar on the overall average and the ALPT listening comprehensive, somewhat less similar on the ALPT reading comprehension. For each of the three tests there is a small average difference between the scores of the two groups favoring the non-CAI students. A part of this results from the 6% higher survival rate of CAI students; since the students who dropped academically were substandard performers, the higher the percentage of drops, the higher the average score of those who remained.

The uncorrected differences between group mean scores are of the size expected to be observed when there is no real difference. The t-values for differences of means of the three tests are 1.2, 1.3, and 0.9, whereas for a difference to be significant at even the .10 level the t-value for 61 degrees of freedom should exceed 1.67.[1] We conclude from the followup tests that there were no major differences in state of learning between CAI and non-CAI students at the end of the Russian ACC course.

---

[1] See, for example, B. J. Winer, Statistical Principles of Experimental Design, McGraw Hill, New York, 1962, pp. 24-31 and p 641.

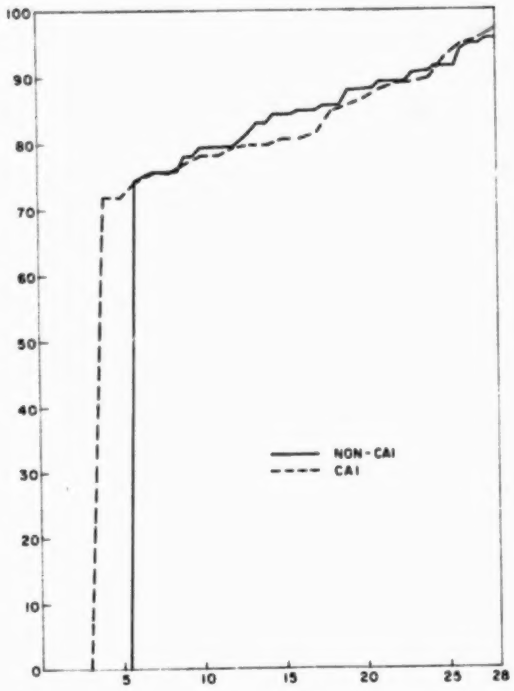


Fig. L-1 - Distribution of Final Averages

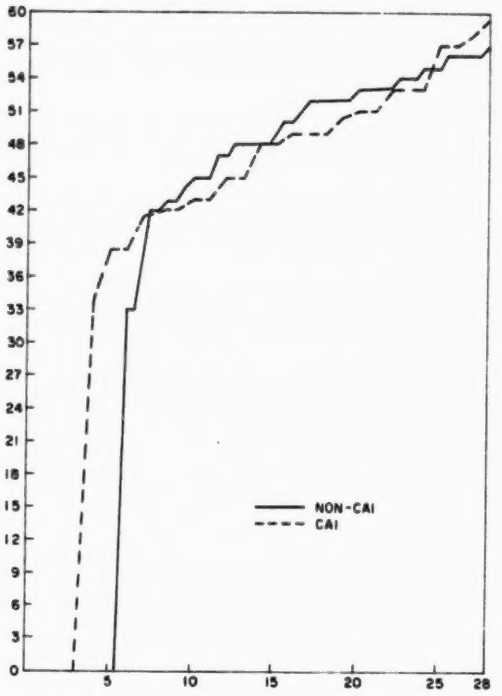


Fig. L-2 - Distribution of ALPT Listening Comprehension

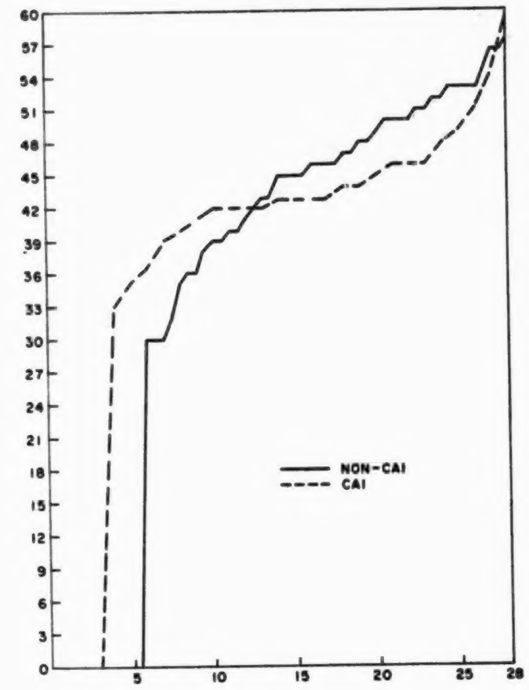


Fig. L-3 - Distribution of ALPT Reading Comprehension

FINAL TEST RESULTS

Student No.	Final Average	ALPT Listening Comp.	ALPT Reading Comp.
1.	administrative drop 4/69		
2.	88	48	49
3.	93	53	48
4.	80	51	43
5.	89	53	42
6.	95	57	51
7.	96	57	54
8.	85	42	43
9.	97	59	59
10.	90	58	46
11.	76	43	33
12.	81	42	45
13.	81	48	44
14.	82	49	39
15.	78	43	43
16.	79	38	43
17.	80	45	42
18.	89	53	46
19.	78	49	42
20.	86	51	46
21.	74	45	44
22.	77	50	41
23.	failed June 9		
24.	87	49	42
25.	72	34	37
26.	failed before April 1		
27.	administrative drop 4/69		
28.	76	41	40
29.	failed before April 1		
30.	72	38	35
31.	96	59	59
32.	90	48	42
33.	96	57	58
34.	95	56	56
35	91	50	40
36.	76	38	36
37.	94	55	50
38.	84	45	45
39.	89	50	42
40.	83	51	38
41.	89	52	41
42.	92	56	49
43.	92	48	53
44.	89	53	52
45.	96	53	54
46.	85	45	45
47.	85	55	47
48.	95	56	56
49.	88	54	50
50.	86	52	46
51.	failed before April 1		
52.	91	53	53
53.	86	52	48
54.	84	48	51
55.	failed before April 1		
56.	75	43	30
57.	88	55	46
58.	83	49	48
59.	81	56	45
60.	91	54	51
61.	74	43	32
62.	88	56	53
63.	76	42	36
64.	failed 9/6/69		
65.	79	47	43
66.	79	45	40
67.	79	42	45
68.	76	48	39
69.	failed before April 1		
70.	78	33	30
71.	failed before April 1		
72.	failed before April 1		
73.	82	52	46
74.	failed before April 1		
75.	84	53	50
76.	failed before April 1		
77.	85	52	47
78.	79	44	35